arXiv:1304.2934v1 [math.PR] 10 Apr 2013

MOD-ϕ CONVERGENCE AND PRECISE DEVIATIONS

VALENTIN FÉRAY, PIERRE-LOÏC MÉLIOT, AND ASHKAN NIKEGHBALI

ABSTRACT. In this paper, we use the framework of mod-ϕ convergence to prove precise large or moderate deviations for quite general sequences of random variables $(X_n)_{n\in\mathbb{N}}$. The random variables considered can be lattice or non-lattice distributed, and single or multi-dimensional; and one obtains precise estimates of the fluctuations $\mathbb{P}[X_n \in t_n B]$, instead of the usual estimates for the rate of exponential decay $\log(\mathbb{P}[X_n \in t_n B])$. The four first sections of the article are devoted to a proof of these abstract results. We then propose new examples covered by this theory and coming from various areas of mathematics: classical probability (multi-dimensional random walks), number theory (statistics of additive arithmetic functions), combinatorics (statistics of random permutations), random matrix theory (characteristic polynomials of random matrices in compact Lie groups), graph theory (number of subgraphs in a random Erdős-Rényi graph), and non-commutative probability (asymptotics of random character values of symmetric groups). In particular, we complete our theory of precise deviations by a concrete method of cumulants and dependency graphs, which applies to many examples of sums of "weakly dependent" random variables.

CONTENTS

Date: April 11, 2013.

1. Introduction

The notion of mod-ϕ convergence has been studied in [JKN11, DKN11, KN10, KN12, BKN13], essentially in connection with problems from number theory, random matrix theory and probability theory. The main idea was to look for a natural renormalization of the characteristic functions of random variables which do not converge in law (instead of a renormalization of the random variables themselves). After this renormalization, the sequence of characteristic functions converges to some non-trivial limiting function. We first recall the definition of mod-ϕ convergence, but in the more restrictive case (but which is the framework for precise large deviations) where the moment generating functions are defined on an open neighborhood of 0.

Definition 1.1. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables, and $\varphi_n(z) = \mathbb{E}[\mathrm{e}^{zX_n}]$ be their moment generating functions, which we assume to all exist in a strip*

$$S_c = \{z,\ -c < \operatorname{Re} z < c\},$$

with c positive real number. We assume that there exists an infinitely divisible distribution ϕ with moment generating function $\int \mathrm{e}^{zx}\phi(dx) \equiv \exp(\eta(z))$, and an analytic function $\psi(z)$ that does not vanish on the real part of S_c, such that locally uniformly in $z \in S_c$,

$$\exp(-t_n\,\eta(z))\,\varphi_n(z) \to \psi(z), \tag{1}$$

where $(t_n)_{n\in\mathbb{N}}$ is some sequence going to $+\infty$. We then say that $(X_n)_{n\in\mathbb{N}}$ converges mod-ϕ, *with parameters $(t_n)_{n\in\mathbb{N}}$ and limiting function ψ. In the following we denote $\psi_n(z)$ the left-hand side of (1).*

Remark 1.2. These hypotheses imply that the X_n's and ϕ have moments of all order; in particular, ϕ cannot be any infinitely divisible distribution (for instance the Cauchy distribution is excluded). In the examples, we shall see that the most common situations are $\phi = \mathcal{N}(0,\sigma^2)$ (centered Gaussian law with variance σ^2) corresponding to the framework of the so-called mod-Gaussian convergence and $\phi = \mathcal{P}(\lambda)$ (Poisson law with parameter λ) corresponding to the case of mod-Poisson convergence. In the sequel we shall use quite often the fact that the local uniform convergence of analytic functions imply those of their derivatives by Cauchy formula.

Remark 1.3. It is to avoid the restrictions mentioned in the previous remark that the theory of mod-ϕ convergence was initially developed with characteristic functions rather than moment generating functions. There are indeed many examples for instance of mod-Cauchy convergence (see e.g. [DKN11, KNN10]) or more generally of mod-ϕ convergence where ϕ is a stable distribution. For lattice examples of situations where the moment generating method would not work the reader can look at [BKN13].

It is immediate to see that mod-ϕ convergence implies a *central limit theorem* if the sequence of parameters t_n goes to infinity (see the remark after Proposition 2.8). But in fact there is much more information encoded in mod-ϕ convergence than merely the central limit theorem. Indeed the works [JKN11, DKN11, KN10, KN12, BKN13] tend to illustrate the fact that mod-ϕ convergence appears as a natural extension of the framework of sums of independent random variables in the sense that many interesting asymptotic results that hold for sums of independent random variables can also be established for a sequence of random variables converging in the mod-ϕ sense. For instance, under some general extra assumptions on the convergence in (1) it is proved in [DKN11] and [KN12]

that one can establish local limit theorems for the random variables X_n. In particular it is shown in [DKN11] that the framework of mod-ϕ convergence extends the classical framework of sums of i.i.d. random variables for local limit theorems (the local limit theorem of Stone appears as a special case of the local limit theorem in the framework of mod-ϕ convergence). On the other hand it applies also to a variety of situations where the random variables under consideration exhibit some dependence structure (e.g. the Riemann zeta function on the critical line, some probabilistic models of primes, the winding number for the planar Brownian motion, the characteristic polynomial of random matrices, finite fields L-functions, etc.). It is also shown in [BKN13] that mod-Poisson convergence (in fact mod-ϕ convergence for ϕ a lattice distribution) implies very sharp distributional approximation in the total variation distance (among other distances) for a large class of random variables. In particular it is shown that the total number of distinct prime divisors $\omega(n)$ of an integer n chosen at random can be approximated in the total variation distance with an arbitrary precision by explicitly computable measures. Besides these quantitative aspects, mod-ϕ convergence also sheds some new light on the nature of some conjectures in analytic number theory. Indeed it is shown in [KN10] that the structure of the limiting function appearing in the moments conjecture for the Riemann zeta function by Keating and Snaith [KS00b] is shared by other arithmetic functions and that the limiting function ψ accounts for the fact that prime numbers do not behave independently of each other. More precisely, the limiting function ψ can be used to measure the deviation of the true result from what the probabilistic models based on a naive independence assumption would predict. One should note that these naive probabilistic models are usually enough to predict central limit theorems for arithmetic functions (e.g. the naive probabilistic model made with a sum of independent Bernoulli random variables to predict the Erdős-Kac central limit theorem for $\omega(n)$ or the stochastic zeta function to predict Selberg's central limit theorem for the Riemann zeta function) but fail to predict accurately mod-ϕ convergence by a factor which is contained in ψ. More generally it seems that the limiting function encodes information about the dependence between the different components of a random vector: indeed it is noted in [KN12] that the log of the characteristic polynomial of a random unitary matrix, as a vector in $\mathbb{R}^2$, converges in the mod-Gaussian sense to a limiting function which is not the product of the limiting functions of each component considered individually although when properly normalized it converges to a Gaussian vector with independent components. But we do not understand yet well the nature of the limiting function in relation with the dependence structure of the random variables. The remark 2.5 of the present paper emphasizes another property of the limiting function ψ: it measures the deviation of the distributions of the X_n's from that of the infinitely divisible distribution ϕ.

More generally the goal of this paper is to prove that the framework of mod-ϕ convergence as described in Definition 1.1 is also suitable to obtain precise (i.e. estimates of the probabilities without the log) large and moderate deviations results for the sequence $(X_n)_{n\in\mathbb{N}}$. As stated in Remark 2.11 our results can be viewed as an extension of the classical results of Bahadur-Rao and Ellis-Gärtner to the more general case described in Definition 1.1, but sometimes with different scales or regimes since we are interested in events of the form $\mathbb{P}[X_n \in t_nB]$ rather than $1/t_n \log \mathbb{P}[X_n \in t_nB]$. In fact we provide an asymptotic expansion in powers of $1/t_n$ whenever the convergence in Definition 1.1 is fast enough. Moreover we are able, at least in the non-lattice case, to establish precise large deviations results in a multidimensional setting. This situation requires more care since the geometry of the Borel set B, when considering $\mathbb{P}[X_n \in t_nB]$, will play a crucial role.

The arguments involved in the proofs of our large deviations results in dimension one are standard but they nonetheless need to be carefully adapted to the framework of Definition 1.1: elementary complex analysis, the method of change of probability measure or tilting due to Cramér or adaptations of Berry-Esseen type inequalities with smoothing techniques. However it should be noted that in the multidimensional case the arguments are more involved and we provide new Berry-Esseen type estimates for Gaussian regular domains or polytopes which have an interest in their own and which extend recent results in [BR10].

Remark 1.4. One should here give much credit to Hwang who has already established some of our results in the lattice case in [Hwa96] using the same framework as in Definition 1.1. The results in [Hwa96] are one dimensional, they do not cover all regimes we shall consider. Moreover lower order terms in the asymptotic expansion for probabilities for large deviations are not considered. The results and ideas of Hwang have already been used by Radziwill in [Rad09] to prove precise large deviations for a large class of additive arithmetic functions. We shall also recover these results in Section 5 as an illustration of the Selberg-Delange method.

These abstract results being established, one needs to provide a large set of (new) examples where these results can be applied. We have thus devoted the last four sections of the paper to examples, from a variety of different areas. Section 5 contains what one would be tempted to call "known examples" in the sense that most of these examples were proved to satisfy mod-ϕ convergence in earlier works. In some of the cases the precise moderate or large deviations results we obtain are new. For instance this is surprisingly the case for the sums of independent random variables and its applications to the loss of symmetry of the multidimensional random walks on $\mathbb{Z}^d$ conditioned to stay far away from the origin. The results of Section 5.5 on the characteristic polynomial of random unitary matrices of different types is also new and can be viewed as completing the previous results by Hughes, Keating and O'Connell [HKO01] on large deviations for the characteristic polynomial as well as the local limit theorems obtained in [KN12, DKN11]. We also recover results of Radziwill [Rad09] on precise large deviations for additive arithmetic functions by carefully recalling the principle of the Selberg-Delange method as well as results by Nikeghbali and Zeindler [NZ13] on precise large deviations for the total number of cycles for random permutations under the general weighted probability measure as an illustration of the singularity analysis method.

We also propose in Section 3.2 some techniques based on control of cumulants (in the mod-Gaussian framework) to establish precise moderate deviations results. Roughly speaking, the idea behind is that we have mod-gaussian convergence when after normalization, the second cumulant explodes, the third one converges and the remaining ones tend to zero. The results of Section 3.2 will be the basis to a whole body of new examples of mod-Gaussian convergence that are introduced from Section 6. More precisely, in Section 6 we combine the results from Section 3.2 with dependency graphs to prove precise moderate deviations for the sum of partially dependent random variables, thus improving and strengthening previous results (all references and details are provided in Section 6). The idea of using bounds on cumulants to show precise deviations for a family of random variable with some given dependency graph is not new, but the bounds we obtain in Theorem 6.2 (and also in Theorem 6.3) are much stronger than those which were previously known. We then apply these results to establish precise moderate deviations

for subgraph counts in Erdős-Rényi random graphs (Theorem 7.1). Eventually, as a last application, in Section 8, we use the machinery of dependency graphs in non-commutative probability spaces, namely, the algebras $\mathbb{C}\mathfrak{S}(n)$ of the symmetric groups, all endowed with the restriction of a trace of the infinite symmetric group $\mathfrak{S}(\infty)$. The technique of cumulants still works and it gives the fluctuations of random integer partitions under so-called central measures in the terminology of Kerov and Vershik. Thus, one obtains a central limit theorem and moderate deviations for the values of the random irreducible characters of symmetric groups under these measures. Moreover, Theorem 8.10 gives an idea of an infinite-dimensional generalization of the notion of mod-Gaussian convergence: the covariance matrix gets replaced by a quadratic form on $\mathbb{C}[x]$, and the limiting function is the exponential of a trilinear form on the same space of polynomials.

We close this introduction by giving a few examples which will guide our intuition throughout the paper. In these examples, it will be useful sometimes to precise the speed of convergence in Definition 1.1. Thus,

Definition 1.5. *We say that the sequence of random variables $(X_n)_{n\in\mathbb{N}}$ converges mod-ϕ at speed $O((t_n)^{-v})$ if the difference of the two sides of Equation (1) can be bounded by $C_K\,(t_n)^{-v}$ for any z in a given compact subset K of S_c. We use the analogue definitions with the $o(\cdot)$ notation.*

Example 1.6. The first example was used in [KNN10] to characterize the set of limiting functions in the setting of mod-ϕ convergence. Let $(Y_n)_{n\in\mathbb{N}}$ be a sequence of centered, independent and identically distributed random variables, with $\mathbb{E}[\mathrm{e}^{zY}] = \mathbb{E}[\mathrm{e}^{zY_1}]$ analytic and non-vanishing on a disc $D(0,R)$, possibly with $R = +\infty$. Set $S_n = Y_1 + \cdots + Y_n$. The cumulant generating series of S_n is

$$\log \mathbb{E}[\mathrm{e}^{zS_n}] = n\,\log \mathbb{E}[\mathrm{e}^{zY}] = n \sum_{r=2}^{\infty} \frac{\kappa^{(r)}(Y)}{r!}\,z^r,$$

which is also analytic on $D(0,R)$. Let $v \geq 3$ be an integer such that $\kappa^{(r)}(Y) = 0$ for $r \in [3, v-1]$, and set $X_n = \frac{S_n}{n^{1/v}}$. One has

$$\log \varphi_n(z) = n^{\frac{v-2}{v}}\,\frac{\kappa^{(2)}(Y)}{2}\,z^2 + \frac{\kappa^{(v)}(Y)}{v!}\,z^v + \sum_{r=v+1}^{\infty} \frac{\kappa^{(r)}(Y)}{r!\,n^{r/v-1}}\,z^r,$$

and locally uniformly on S_R the right-most term is bounded by $\frac{C}{n^{1/v}}$, so it converges locally uniformly to 0. Consequently,

$$\psi_n(z) = \exp\left(-n^{\frac{v-2}{v}}\,\frac{\sigma^2 z^2}{2}\right)\,\varphi_n(z) \to \exp\left(\frac{\kappa^{(v)}(Y)}{v!}\,z^v\right),$$

that is, $(X_n)_{n\in\mathbb{N}}$ converges in the mod-Gaussian sense with parameters $t_n = \sigma^2\,n^{\frac{v-2}{v}}$, speed $O(n^{-1/v})$ and limiting function $\psi(z) = \exp(\kappa^{(v)}(Y)\,z^v/v!)$.

Example 1.7. More generally, let $(S_n)_{n\in\mathbb{N}}$ be a sequence of real-valued centered random variables that admit moments of all order, and such that $|\kappa^{(r)}(S_n)| \leq (Cr)^r\,n^{a+br}$ for all $r \geq 2$ and for some constants C, a, b. Assume moreover that there exists an integer $v \geq 3$ such that

(1) $\kappa^{(r)}(S_n) = 0$ for all $3 \leq r < v$ and all $n \in \mathbb{N}$;

(2) $\lim_{n\to\infty} \frac{\kappa^{(2)}(S_n)}{n^{a+2b}} = \sigma^2$ and $\lim_{n\to\infty} \frac{\kappa^{(v)}(S_n)}{n^{a+vb}} = L$ exist; more precisely,

$$\kappa^{(2)}(S_n) = \sigma^2\, n^{a+2b}\, \left(1 + o\left(n^{-a(\frac{v-2}{v})}\right)\right) \qquad ; \qquad \kappa^{(v)}(S_n) = L\, n^{a+bv}\, \big(1+o(1)\big).$$

Set $X_n = \frac{S_n}{n^{\frac{a}{v}+b}}$. The cumulant generating series of X_n is

$$\begin{aligned}\log \varphi_n(z) &= \frac{\kappa^{(2)}(S_n)}{2\, n^{\frac{2a}{v}+2b}}\, z^2 + \frac{\kappa^{(v)}(S_n)}{v!\, n^{a+vb}}\, z^v + \sum_{r=v+1}^{\infty} \frac{\kappa^{(r)}(S_n)}{r!\, n^{\frac{r}{v}a+rb}}\, z^r \\ &= \frac{\sigma^2}{2}\, n^{a\,\frac{v-2}{2}}\, z^2 + \frac{L}{v!}\, z^v + \sum_{r=v+1}^{\infty} \frac{\kappa^{(r)}(S_n)}{r!\, n^{\frac{r}{v}a+rb}}\, z^r + o(1),\end{aligned}$$

where the $o(1)$ is locally uniform. The remaining series is locally uniformly bounded in absolute value by

$$\sum_{r=v+1}^{\infty} C^r\, \frac{r^r}{r!}\, \frac{1}{n^{a\frac{r-v}{v}}}\, R^r \leq n^a \sum_{r=v+1}^{\infty} \left(\frac{\mathrm{e}\, CR}{n^{\frac{a}{v}}}\right)^r = n^{-\frac{a}{v}}\, \frac{(\mathrm{e}\, CR)^{v+1}}{1 - \mathrm{e}\, CR\, n^{-\frac{a}{v}}} \to 0.$$

Hence,

$$\psi_n(z) = \exp\left(-n^{a\,\frac{v-2}{v}}\, \frac{\sigma^2 z^2}{2}\right)\, \varphi_n(z) \to \exp\left(\frac{L}{v!}\, z^v\right)$$

locally uniformly on $\mathbb{C}$, so one has again mod-Gaussian convergence, with parameters $t_n = \sigma^2\, n^{a\frac{v-2}{v}}$ and limiting function $\psi(z) = \mathrm{e}^{\frac{L}{v!}\, z^v}$. In the case of sums of i.i.d. variables, $a = 1$ and $b = 0$. However, this framework includes many more examples than sums of i.i.d. variables: in particular, in Section 6, we show that such bounds on cumulants can be obtained for sums of partially dependent random variables.

Example 1.8. Denote X_n the number of disjoint cycles (including fixed points) of a random permutation chosen uniformly in the symmetric group $\mathfrak{S}_n$. Feller's coupling (*cf.* [ABT03, Chapter 1]) shows that $X_n =_{(\mathrm{law})} \sum_{i=1}^n \mathcal{B}_{(1/i)}$, where $\mathcal{B}_p$ denotes a Bernoulli variable equal to 1 with probability p and to 0 with probability $1-p$, and the Bernoulli variables are independent in the previous expansion. So,

$$\mathbb{E}[\mathrm{e}^{zX_n}] = \prod_{i=1}^{n} \left(1 + \frac{\mathrm{e}^z - 1}{i}\right) = \mathrm{e}^{H_n(\mathrm{e}^z-1)} \prod_{i=1}^{n} \frac{1 + \frac{\mathrm{e}^z-1}{i}}{\mathrm{e}^{\frac{\mathrm{e}^z-1}{i}}}$$

where $H_n = \sum_{i=1}^n \frac{1}{i} = \log n + \gamma + O\big(\frac{1}{n}\big)$. The product in the right-hand side converges locally uniformly to an entire function, therefore,

$$\mathbb{E}[\mathrm{e}^{zX_n}]\, \mathrm{e}^{-(\mathrm{e}^z-1)\log n} \to \mathrm{e}^{\gamma\,(\mathrm{e}^z-1)} \prod_{i=1}^{\infty} \frac{1 + \frac{\mathrm{e}^z-1}{i}}{\mathrm{e}^{\frac{\mathrm{e}^z-1}{i}}} = \frac{1}{\Gamma(\mathrm{e}^z)}$$

locally uniformly, *i.e.*, one has mod-Poisson convergence with parameters $t_n = \log n$ and limiting function $\frac{1}{\Gamma(\mathrm{e}^z)}$. Moreover, the speed of convergence is a $O\big(\frac{1}{n}\big)$, hence, a $o((t_n)^{-v})$ for any integer v.

Acknowledgements

The authors would like to thank Martin Wahl for his input at the beginning of this project and for sharing with us his ideas. We would also like to address special thanks to Andrew Barbour, Reda Chhaibi and Kenny Maples for many fruitful discussions which helped us improve some of our arguments.

2. Large deviations in the case of lattice distributions

In the theory of large deviations, the following transform on (convex) functions is known to play a central role (see *e.g.* [DZ98, §2.2]):

Definition 2.1. *The Legendre-Fenchel transform of a function η is defined by:*

$$F(x) = \sup_{h\in\mathbb{R}} (hx - \eta(h)).$$

This is an involution on convex lower semi-continuous functions.

If η is the logarithm of the moment generating series of a random variable, then F is always non-negative, and the unique h maximizing $hx - \eta(h)$, if it exists, is then defined by the implicit equation $\eta'(h(x)) = x$. This implies the following useful identities:

$$F(x) = xh(x) - \eta(h(x)) \qquad ; \qquad F'(x) = h(x) \qquad ; \qquad F''(x) = \frac{1}{\eta''(h(x))}.$$

Example 2.2. If $\eta(z) = mz + \frac{\sigma^2 z^2}{2}$ (Gaussian variable with mean m and variance σ^2), then

$$h(x) = \frac{x-m}{\sigma^2} \qquad ; \qquad F_{\mathcal{N}(m,\sigma^2)}(x) = \frac{(x-m)^2}{2\sigma^2}$$

whereas if $\eta(z) = \lambda(\mathrm{e}^z - 1)$ (Poisson law with parameter λ), then

$$h(x) = \log\frac{x}{\lambda} \qquad ; \qquad F_{\mathcal{P}(\lambda)}(x) = \begin{cases} x\log\frac{x}{\lambda} - (x-\lambda) & \text{if } x > 0, \\ +\infty & \text{otherwise} \end{cases}.$$

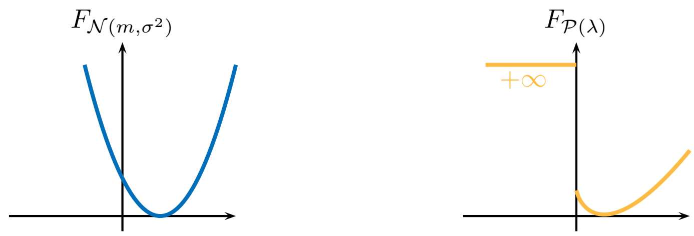


Figure 1. The Legendre-Fenchel transforms of a Gaussian law and of a Poisson law.

2.1. **Large deviations in the scale $O(t_n)$.** In this section we suppose that the X_n's and the infinitely divisible distribution ϕ are lattice distributed, and that ϕ has minimal lattice $\mathbb{Z}$. This means that the X_n's take values in $\mathbb{Z}$, and that the characteristic function

$$\widehat{\phi}(u) = \phi(\mathrm{e}^{\mathrm{i}u}) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}u}\,\phi(du) = \exp(\eta(\mathrm{i}u))$$

of the distribution ϕ has minimal period 2π. In particular, for every $u \in (0, 2\pi)$, $|\exp(\eta(\mathrm{i}u))| < 1$, since the smallest period of the characteristic function of a $\mathbb{Z}$-valued infinitely divisible distribution is also the smallest $u > 0$ such that $|\phi(\mathrm{e}^{\mathrm{i}u})| = 1$.

Lemma 2.3. *Let X be a $\mathbb{Z}$-valued random variable whose generating function $\varphi_X(z) = \mathbb{E}[\mathrm{e}^{zX}]$ converges absolutely in the strip S_c. For $k \in \mathbb{Z}$,*

$$\forall h \in (-c,c),\ \ \mathbb{P}[X=k] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-k(h+\mathrm{i}u)}\,\varphi_X(h+\mathrm{i}u)\,du;$$

$$\forall h \in (0,c),\ \ \mathbb{P}[X\geq k] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\mathrm{e}^{-k(h+\mathrm{i}u)}}{1-\mathrm{e}^{-(h+\mathrm{i}u)}}\,\varphi_X(h+\mathrm{i}u)\,du.$$

Proof. Since

$$\varphi_X(h+\mathrm{i}u) = \sum_{k\in\mathbb{Z}} \mathbb{P}[X=k]\,\mathrm{e}^{k(h+\mathrm{i}u)},$$

$\mathbb{P}[X=k]\,\mathrm{e}^{kh}$ is the k-th Fourier coefficient of the 2π-periodic and smooth function $u \mapsto \varphi_X(h+\mathrm{i}u)$; this leads to the first formula. Then, assuming also $h>0$,

$$\mathbb{P}[X\geq k] = \sum_{l=k}^{\infty} \mathbb{P}[X=l] = \sum_{l=k}^{\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-l(h+\mathrm{i}u)}\,\varphi_X(h+\mathrm{i}u)\,du,$$

and the sum of the moduli of the functions on the right-hand side is dominated by the integrable function $\frac{\mathrm{e}^{-kh}}{1-\mathrm{e}^{-h}}\,\varphi_X(h)$; so by Lebesgue's dominated convergence theorem, one can exchange the integral and the series, which yields the second equation. $\square$

We now work under the assumptions of Definition 1.1, and furthermore, we assume that the convergence is at speed $O((t_n)^{-v})$.

Theorem 2.4. *Let x be a real number in the range of $\eta'_{|(-c,c)}$, h defined by the implicit equation $\eta'(h) = x$. We assume $t_n x \in \mathbb{N}$.*

(1) The following expansion holds:

$$\begin{aligned}\mathbb{P}[X_n = t_n x] &= \frac{\exp(-t_n F(x))}{\sqrt{2\pi t_n \eta''(h)}}\left(\psi(h) + \frac{a_1}{t_n} + \frac{a_2}{(t_n)^2} + \cdots + \frac{a_{v-1}}{(t_n)^{v-1}} + O\!\left(\frac{1}{(t_n)^v}\right)\right)\\ &= \exp(-t_n F(x))\sqrt{\frac{F''(x)}{2\pi t_n}}\left(\psi(F'(x)) + \frac{a_1}{t_n} + \cdots + \frac{a_{v-1}}{(t_n)^{v-1}} + O\!\left(\frac{1}{(t_n)^v}\right)\right).\end{aligned}$$

The a_k's are rational fractions in the derivatives of η and ψ at h. More precisely, denote

$$\Delta_n(w) = t_n\left(\eta\left(h + \frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}\right) - \eta'(h)\,\frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}} + \frac{w^2}{2t_n}\right)$$

$$f_n(w) = \psi\left(h + \frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}\right)\exp(\Delta_n(w)) = \sum_{k=0}^{\infty} \frac{\alpha_k(w)}{(t_n)^{k/2}},$$

the last expansion holding in a neighborhood of zero. The coefficient $\alpha_{2k}(w)$ is an even polynomial in w with valuation $2k$ and coefficients which are polynomials in the derivatives of ψ and η at h, and in $\frac{1}{\eta''(h)}$. Then,

$$a_k = \int_{\mathbb{R}} \alpha_{2k}(w)\,\frac{\mathrm{e}^{-\frac{w^2}{2}}}{\sqrt{2\pi}}\,dw,$$

and in particular,

$$\begin{aligned} a_0 &= \psi(h); \\ a_1 &= -\frac{1}{2}\,\frac{\psi''(h)}{\eta''(h)} + \frac{1}{24}\,\frac{\psi(h)\,\eta^{(4)}(h) + 4\,\psi'(h)\,\eta^{(3)}(h)}{(\eta''(h))^2} - \frac{15}{72}\,\frac{\psi(h)\,(\eta^{(3)}(h))^2}{(\eta''(h))^3}. \end{aligned}$$

(2) Similarly, if x is a real number in the range of $\eta'_{|(0,c)}$, then

$$\mathbb{P}[X_n \geq t_n x] = \frac{\exp(-t_n F(x))}{\sqrt{2\pi t_n \eta''(h)}}\,\frac{1}{1-\mathrm{e}^{-h}}\left(\psi(h) + \frac{b_1}{(t_n)} + \cdots + \frac{b_{v-1}}{(t_n)^{v-1}} + O\!\left(\frac{1}{(t_n)^v}\right)\right),$$

where the b_k's are obtained by the same recipe as the a_k's, but starting from the power series

$$g_n(w) = \frac{1-\exp(-h)}{1-\exp\left(-h - \frac{\mathrm{i}w}{\sqrt{t_n \eta''(h)}}\right)}\,f_n(w).$$

Remark 2.5. For $x > \eta'(0)$, the first term of the expansion

$$\frac{\exp(-t_n F(x))}{\sqrt{2\pi t_n \eta''(h)}}$$

is also the leading term in the asymptotics of $\mathbb{P}[Y_{t_n} = t_n x]$, where $(Y_t)_{t\in\mathbb{R}_+}$ is the Lévy process associated to the analytic function $\eta(z)$. Thus, the $\psi(h)$ measures the difference between the distribution of X_n and the distribution of Y_{t_n} in the interval $(t_n\eta'(0), t_n\eta'(c))$.

Remark 2.6. If the convergence is faster than any negative power of t_n, then one can simplify the statement of the theorem as follows: as formal power series in t_n,

$$\sqrt{2\pi t_n \eta''(h)}\,\exp(t_n F(x))\,\mathbb{P}[X_n = t_n x] = \int_{\mathbb{R}} f_n(w)\,\mathrm{e}^{-\frac{w^2}{2}}\,dw,$$

i.e., the expansions of both sides up to any given power $O\!\left(\frac{1}{(t_n)^v}\right)$ agree.

Remark 2.7. If one wants to remove the condition $t_n x \in \mathbb{N}$, then one can only keep the first term of the expansion, since $t_n x - \lfloor t_n x \rfloor = O(1)$ in general. Thus, the first-order expansions of Theorem 2.4 hold for any real number $x \in (\eta'(0), \eta'(c))$, with a remainder that is a $O(\frac{1}{\sqrt{t_n}})$ uniform on compact subsets of $(\eta'(0), \eta'(c))$.

Proof. With the notations of Definition 1.1, the first equation of Lemma 2.3 becomes

$$\begin{aligned} \mathbb{P}[X_n = t_n x] &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-t_n x(h+\mathrm{i}u)}\,\varphi_n(h+\mathrm{i}u)\,du \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-t_n x h}\,\mathrm{e}^{t_n(\eta(h+\mathrm{i}u)-\mathrm{i}ux)}\,\psi_n(h+\mathrm{i}u)\,du \\ &= \frac{\mathrm{e}^{-t_n F(x)}}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{t_n(\eta(h+\mathrm{i}u)-\eta(h)-\mathrm{i}u\eta'(h))}\,\psi_n(h+\mathrm{i}u)\,du. \qquad (2) \end{aligned}$$

We perform the Laplace method on (2), and to this purpose we split the integral in two parts. Fix $\delta > 0$, and denote $q_\delta = \max_{u \in (-\pi,\pi)\setminus(-\delta,\delta)} |\exp(\eta(h+\mathrm{i}u)-\eta(h))|$. This is strictly smaller than 1 (*cf.* Lemma 2.10 hereafter), since

$$\exp(\eta(h+\mathrm{i}u)-\eta(h)) = \frac{\mathbb{E}[\mathrm{e}^{(h+\mathrm{i}u)X}]}{\mathbb{E}[\mathrm{e}^{hX}]} = \mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{\mathrm{i}uX}]$$

is the characteristic function of X under the new probability $d\mathbb{Q}(\omega) = \frac{\mathrm{e}^{hX(\omega)}}{\mathbb{E}[\mathrm{e}^{hX}]} d\mathbb{P}(\omega)$.

As a consequence, if $I_{(-\delta,\delta)}$ and $I_{(-\delta,\delta)^c}$ denote the two parts of (2) corresponding to $\int_{-\delta}^{\delta}$ and $\int_{-\pi}^{-\delta} + \int_{\delta}^{\pi}$, then

$$I_{(-\delta,\delta)^c} \leq \frac{\mathrm{e}^{-t_n F(x)}}{2\pi} \int_{(-\delta,\delta)^c} (q_\delta)^{t_n} |\psi_n(h+\mathrm{i}u)|\,du \leq 2\,(\mathrm{e}^{-F(x)}\,q_\delta)^{t_n} \max_{u\in(-\pi,\pi)} |\psi(h+\mathrm{i}u)|$$

for n big enough, since ψ_n converges uniformly towards ψ on the compact set $K = h + \mathrm{i}[-\pi,\pi]$. Since $q_\delta < 1$, for any $\delta > 0$ fixed, $I_{(-\delta,\delta)^c}\,\mathrm{e}^{t_n F(x)}$ goes to 0 faster than any negative power of t_n, so $I_{(-\delta,\delta)^c}$ is negligible in the asymptotics.

As for the other part, we can first replace ψ_n by ψ up to a $(1+O((t_n)^{-v}))$, since the integral is taken on a compact subset of S_c. We then set $u = \frac{w}{\sqrt{t_n \eta''(h)}}$:

$$I_{(-c,c)} = \left(1 + O\left(\frac{1}{(t_n)^v}\right)\right) \frac{\mathrm{e}^{-t_n F(x)}}{2\pi\sqrt{t_n\eta''(h)}} \int_{-\delta\sqrt{t_n\eta''(h)}}^{\delta\sqrt{t_n\eta''(h)}} \psi\left(h + \frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}\right) \mathrm{e}^{t_n\Delta(w)}\,\mathrm{e}^{-\frac{w^2}{2}}\,dw, \tag{3}$$

where $\Delta(w)$ is the Taylor expansion

$$\begin{aligned}
\eta\,(h+\mathrm{i}u) - \eta(h) - \eta'(h)\,(\mathrm{i}u) - \frac{\eta''(h)}{2}\,(\mathrm{i}u)^2 &= \sum_{k=3}^{2v+1} \frac{\eta^{(k)}(h)}{k!} \left(\frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}\right)^k + O\left(\frac{1}{(t_n)^{v+1}}\right) \\
&= \frac{1}{t_n}\left(-\frac{w^2}{\eta''(h)} \sum_{k=1}^{2v-1} \frac{\eta^{(k+2)}(h)}{(k+2)!} \left(\frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}\right)^k + O\left(\frac{1}{(t_n)^v}\right)\right).
\end{aligned}$$

We also replace ψ by its Taylor expansion

$$\psi\left(h + \frac{\mathrm{i}v}{\sqrt{t_n\eta''(h)}}\right) = \sum_{k=0}^{2v-1} \frac{\psi^{(k)}(h)}{k!} \left(\frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}\right)^k + O\left(\frac{1}{(t_n)^v}\right).$$

Thus, if one sets

$$\begin{aligned}
f_n(w) &= \left(\sum_{k=0}^{2v-1} \frac{\psi^{(k)}(h)}{k!} \left(\frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}\right)^k\right) \exp\left(-\frac{w^2}{\eta''(h)} \sum_{k=1}^{2v-1} \frac{\eta^{(k+2)}(h)}{(k+2)!} \left(\frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}\right)^k\right) \\
&= \sum_{k=0}^{2v-1} \frac{\alpha_k(w)}{(t_n)^{k/2}} + O\left(\frac{1}{(t_n)^v}\right),
\end{aligned}$$

then one can replace $\psi_n(h+\mathrm{i}u)\,\mathrm{e}^{t_n\Delta(w)}$ by $f_n(w)$ in Equation (3), and moreover, each coefficient $\alpha_k(w)$ writes as

$$\alpha_k(w) = \alpha_{k,0}(h)\left(\frac{w}{\sqrt{\eta''(h)}}\right)^k + \alpha_{k,1}(h)\left(\frac{w}{\sqrt{\eta''(h)}}\right)^{k+2} + \cdots + \alpha_{k,r}(h)\left(\frac{w}{\sqrt{\eta''(h)}}\right)^{k+2r}$$

with the $\alpha_{k,r}(h)$'s polynomials in the derivatives of ψ and η at point h. So,

$$I_{(-c,c)} = \left(1 + O\left(\frac{1}{(t_n)^v}\right)\right) \frac{\mathrm{e}^{-t_n F(x)}}{\sqrt{2\pi t_n \eta''(h)}} \left(\sum_{k=0}^{2v-1} \int_{-\delta\sqrt{t_n\eta''(h)}}^{\delta\sqrt{t_n\eta''(h)}} \frac{\alpha_k(w)}{(t_n)^{k/2}} \frac{\mathrm{e}^{-\frac{w^2}{2}}}{\sqrt{2\pi}}\, dw\right).$$

For any power w^m,

$$\left|\int_{-\infty}^{\infty} w^m \frac{\mathrm{e}^{-\frac{w^2}{2}}}{\sqrt{2\pi}}\, dw - \int_{-\delta\sqrt{t_n\eta''(h)}}^{\delta\sqrt{t_n\eta''(h)}} w^m \frac{\mathrm{e}^{-\frac{w^2}{2}}}{\sqrt{2\pi}}\, dw\right|$$

is smaller than any negative power of t_n as n goes to infinity: indeed, by integration by parts, one can expand the difference as $\mathrm{e}^{-\delta^2 t_n \eta''(h)/2} R_m(\sqrt{t_n})$, where R_m is a rational fraction that depends on m, h, δ and on the order of the expansion needed. Therefore, one can take the full integrals in the previous formula. On the other hand, the odd moments of the Gaussian distribution vanish. One concludes that

$$\mathbb{P}[X_n = t_n x] = \frac{\mathrm{e}^{-t_n F(x)}}{\sqrt{2\pi t_n \eta''(h)}} \left(\sum_{k=0}^{v-1} \frac{1}{(t_n)^k} \left(\int_{\mathbb{R}} \alpha_{2k}(w) \frac{\mathrm{e}^{-\frac{w^2}{2}}}{\sqrt{2\pi}}\, dw\right) + O\left(\frac{1}{(t_n)^v}\right)\right),$$

and each integral $\int_{\mathbb{R}} \alpha_{2k}(w) \frac{\mathrm{e}^{-\frac{w^2}{2}}}{\sqrt{2\pi}}\, dw$ is equal to

$$\frac{\alpha_{2k,0}(h)\,(2k-1)!!}{(\eta''(h))^k} + \cdots + \frac{\alpha_{2k,r}(h)\,(2k+2r-1)!!}{(\eta''(h))^{k+r}}$$

where $(2m-1)!!$ is the double factorial $(2m-1)(2m-3)\cdots 3\,1$, that is to say the $2m$-th moment of the Gaussian distribution. This ends the proof of the first part of our Theorem, the second formula coming from the identities $h = F'(x)$ and $\eta''(h) = \frac{1}{F''(x)}$. The second part is exactly the same, up to the factor

$$\frac{1}{1-\mathrm{e}^{-h-\mathrm{i}u}} = \frac{1}{1-\mathrm{e}^{-h}} \left(\frac{1-\mathrm{e}^{-h}}{1-\mathrm{e}^{-h-\frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}}}\right)$$

in the integrals. $\square$

2.2. Central limit theorem at the scales $o(t_n)$ and $o((t_n)^{2/3})$.

By making an expansion around $\eta'(0)$ and modifying a little the arguments of the previous proof, one also gets the following central limit theorem:

Proposition 2.8. *Assume $y = o((t_n)^{1/6})$. Then, under the assumptions of Theorem 2.4,*

$$\mathbb{P}\left[X_n \geq t_n\eta'(0) + \sqrt{t_n\eta''(0)}\, y\right] = \mathbb{P}[\mathcal{N}(0,1) \geq y]\,(1+o(1))\,.$$

More generally, if $s = o(1)$, $x = \eta'(0) + s$ and h is the solution of $\eta'(h) = x$, then

$$\mathbb{P}[X_n \geq t_n(\eta'(0)+s)] = \mathrm{e}^{t_n\left(\frac{\eta''(h)\,h^2}{2} - F(\eta'(0)+s)\right)}\, \mathbb{P}\left[\mathcal{N}(0,1) \geq h\sqrt{t_n\,\eta''(h)}\right]\,\big(1+o(1)\big).$$

Remark 2.9. In the case $y = O(1)$, which is the classical central limit, this follows immediately from the assumptions of Definition 1.1, since by a Taylor expansion around 0 of η the characteristic functions of the rescaled r.v.

$$Y_n = \frac{X_n - t_n\eta'(0)}{\sqrt{t_n\eta''(0)}}$$

converge pointwise to $\mathrm{e}^{-\frac{\zeta^2}{2}}$, the characteristic function of the standard Gaussian distribution. The improvement here is the weaker assumption $y = o((t_n)^{1/6})$.

Lemma 2.10. *There exists a constant $D > 0$ only depending on η, and an interval $(-\varepsilon, \varepsilon)$, such that for all $h \in (-\varepsilon, \varepsilon)$ and all δ small enough,*

$$q_\delta = \max_{u \in (-\delta,\delta)^c} |\exp(\eta(h+\mathrm{i}u) - \eta(h))| \leq 1 - D\,\delta^2.$$

Proof. We have seen that $\exp(\eta(h+\mathrm{i}u) - \eta(h))$ can be interpreted as the characteristic function of X under the new probability measure $d\mathbb{Q} = \frac{\mathrm{e}^{hX}}{\mathbb{E}[\mathrm{e}^{hX}]}\,d\mathbb{P}$. So, for any u,

$$\begin{aligned}|\exp(\eta(h+\mathrm{i}u) - \eta(h))|^2 = \left|\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{\mathrm{i}uX}]\right|^2 &= \sum_{n,m \in \mathbb{Z}} \mathbb{Q}[X=n]\,\mathbb{Q}[X=m]\,\mathrm{e}^{\mathrm{i}u(n-m)} \\ &= \sum_{k\in\mathbb{Z}} \left(\sum_{n-m=k} \mathbb{Q}[X=n]\,\mathbb{Q}[X=m]\right) \cos ku.\end{aligned}$$

Since $\mathbb{Z}$ is the minimal lattice for ϕ, there exists n, m with $n - m = 1$ and $\mathbb{P}[X = n] \neq 0$, $\mathbb{P}[X = m] \neq 0$. Then one also has $\mathbb{Q}[X = n] \neq 0$, $\mathbb{Q}[X = m] \neq 0$, and

$$\mathbb{Q}[X=n]\,\mathbb{Q}[X=m] \geq 15\,D > 0$$

for h small enough. As $\cos u \leq 1 - \frac{u^2}{5}$ for all $u \in (-\pi, \pi)$,

$$\begin{aligned}&|\exp(\eta(h+\mathrm{i}u) - \eta(h))|^2 \leq 1 + 15\,D\,(\cos u - 1) \leq 1 - 3\,D\,u^2;\\ &q_\delta \leq \sqrt{1 - 3\,D\,\delta^2} \leq 1 - D\,\delta^2 \text{ for } \delta \text{ small enough.}\end{aligned}$$

We also refer to [Ess45, Theorem 6] for a general result on the Lebesgue measure of the set of points such that the characteristic function of a distribution is bigger in absolute value than $1 - \delta^2$. ☐

Proof of Proposition 2.8. Notice that $\eta''(0) \neq 0$ since this is the variance of the law ϕ, which is non-zero because of the hypotheses put on its characteristic function. Set $t_n x = t_n(\eta'(0) + s)$, and assume $s = o(1)$. The analogue of Equation (2) reads in our setting

$$\mathbb{P}[X_n \geq t_n x] = \frac{\mathrm{e}^{-t_n F(x)}}{2\pi} \int_{-\pi}^{\pi} \frac{\mathrm{e}^{t_n(\eta(h+\mathrm{i}u) - \eta(h) - \mathrm{i}u\eta'(h))}}{1 - \mathrm{e}^{-h-\mathrm{i}u}}\,\psi_n(h+\mathrm{i}u)\,du. \tag{4}$$

Since $h'(x) = F''(x) = \frac{1}{\eta''(x)}$, $h = \frac{s}{\eta''(0)} + O(s^2)$. The same argument as in the proof of Theorem 2.4 shows that the integral over $(-\delta, \delta)^c$ is bounded by $C\,\delta\,(q_\delta)^{t_n}$, where $q_\delta < 1$, and $C\,\delta$ (with C a constant independent from s and δ) comes from the computation of

$$\max_{u \in (-\delta,\delta)^c} \left|\frac{\psi(h+\mathrm{i}u)}{1 - \mathrm{e}^{-h-\mathrm{i}u}}\right|.$$

In the following we shall need to make δ go to zero sufficiently fast, but with $\delta\sqrt{t_n \eta''(0)}$ still going to infinity. Thus, set $\delta = (t_n)^{-2/5}$, so that in particular $(t_n)^{-1/2} \ll \delta \ll (t_n)^{-1/3}$. Notice that $I_{(-c,c)}\,\mathrm{e}^{t_n F(x)}$ still goes to zero faster than any power of t_n; indeed,

$$(q_\delta)^{t_n} \leq \left(1 - \frac{D}{(t_n)^{4/5}}\right)^{t_n} \leq \mathrm{e}^{-D\,(t_n)^{1/5}}.$$

The other part of (4) is

$$\frac{\mathrm{e}^{-t_n F(x)}}{2\pi\sqrt{t_n\eta''(h)}} \int_{-\delta\sqrt{t_n\eta''(h)}}^{\delta\sqrt{t_n\eta''(h)}} \psi\left(h+\frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}\right)\,\mathrm{e}^{t_n\Delta(w)}\,\frac{\mathrm{e}^{-\frac{w^2}{2}}}{1-\mathrm{e}^{-h-\frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}}}\,dw,$$

up to a factor $(1+o(1))$. Let us analyze each part of the integral:

- The difference between $\psi\left(h+\frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}\right)$ and $\psi(0)$ is bounded by
$$\max_{z\in[-s,s]+\mathrm{i}[-\delta,\delta]} |\psi(z)-\psi(0)| = o(1)$$
by continuity of ψ, so one can replace the term with ψ by the constant $\psi(0)=1$, up to factor $(1+o(1))$.
- The term $\Delta(w)$ has for Taylor expansion
$$\frac{\eta^{(3)}(h)}{6}\left(\frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}\right)^3 + O\left(\frac{1}{(t_n)^2}\right),$$
so $t_n\,\Delta(w)$ is bounded by a $O(t_n\,\delta^3)$, which is a $o(1)$ since $\delta \ll (t_n)^{-1/3}$. So again one can replace $\mathrm{e}^{t_n\Delta(w)}$ by the constant 1.
- The Taylor expansion of $\left(1-\mathrm{e}^{-h-\frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}}\right)^{-1}$ is $\frac{1}{h+\frac{\mathrm{i}w}{\sqrt{t_n\eta''(h)}}}(1+o(1))$. Hence,
$$\begin{aligned}\mathbb{P}\left[X_n \geq t_n(\eta'(0)+s)\right] &= \frac{\mathrm{e}^{-t_nF(x)}}{2\pi}\left(\int_{\mathbb{R}} \frac{\mathrm{e}^{-\frac{w^2}{2}}}{\sqrt{t_n\eta''(h)}\,h+\mathrm{i}w}\,dw\right)\big(1+o(1)\big)\\ &= \mathrm{e}^{-t_nF(x)+\frac{h^2\,t_n\,\eta''(h)}{2}}\,\mathbb{P}\Big[\mathcal{N}(0,1)\geq h\,\sqrt{t_n\eta''(h)}\Big]\,\big(1+o(1)\big).\end{aligned}$$
Indeed, setting $\beta = h\sqrt{t_n\,\eta''(h)}$, one has:
$$\begin{aligned}\mathrm{e}^{-\frac{\beta^2}{2}}\,\frac{1}{2\pi}\int_{\mathbb{R}}\frac{\mathrm{e}^{-\frac{w^2}{2}}}{\beta+\mathrm{i}w}\,dw &= \frac{1}{2\mathrm{i}\pi}\oint_{\Gamma=\beta+\mathrm{i}\mathbb{R}} \frac{\mathrm{e}^{\frac{(z-\beta)^2-\beta^2}{2}}}{z}\,dz\\ &= \int_\beta^\infty \left(\frac{1}{2\mathrm{i}\pi}\oint_{\Gamma=\alpha+\mathrm{i}\mathbb{R}} \mathrm{e}^{\frac{(z-\alpha)^2-\alpha^2}{2}}\,dz\right)d\alpha\\ &= \int_\beta^\infty \frac{\mathrm{e}^{-\frac{\alpha^2}{2}}}{\sqrt{2\pi}}\,d\alpha = \mathbb{P}[\mathcal{N}(0,1)\geq\beta]\end{aligned}$$
by using a calculus of residue on the second line to differentiate with respect to β.

Hence, we have shown so far:

$$\mathbb{P}\left[X_n \geq t_n(\eta'(0)+s)\right] = \mathrm{e}^{-t_nF(\eta'(0)+s)}\,\mathrm{e}^{\frac{\beta^2}{2}}\,\mathbb{P}[\mathcal{N}(0,1)\geq\beta]\big(1+o(1)\big). \tag{5}$$

Set $y = s\sqrt{t_n/\eta''(0)}$, and suppose that $y = o((t_n)^{1/6})$, or, equivalently, $s = o((t_n)^{-1/3})$ — this is stronger than the previous assumption $s = o(1)$. Let us then see how everything is transformed.

- By making a Taylor expansion around $\eta'(0)$ of the Legendre-Fenchel transform, we get
$$F(x) = F(\eta'(0)) + F'(\eta'(0))\, s + \frac{F''(\eta'(0))}{2}\, s^2 + O(s^3) = \frac{y^2}{2t_n} + o(t_n^{-1}),$$
so $\mathrm{e}^{-t_n F(\eta'(0)+s)} \simeq \mathrm{e}^{-\frac{y^2}{2}}$.
- On the other hand,
$$\beta = h\sqrt{t_n\, \eta''(h)} = \frac{s}{\eta''(0)}\,(1+O(s))\,\sqrt{t_n\,(\eta''(0)+O(s))} = y\,(1+O(s)) = y\,(1+o((t_n)^{-1/3}))$$
Consequently, $\beta^2 = y^2(1+o((t_n)^{-1/3})) = y^2 + o(1)$, so $\mathrm{e}^{\frac{\beta^2}{2}}$ can be replaced safely by $\mathrm{e}^{\frac{y^2}{2}}$, which compensates the previous term.
- Finally, fix y, and denote $F_y(\lambda) = \mathbb{P}[\mathcal{N}(0,1) \geq \lambda y]$. Then, for $|\lambda|$ say between $\frac{1}{2}$ and 2,
$$|F_y'(\lambda)| = \left|\frac{y}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{\lambda^2 y^2}{2}}\right| \leq \max_{y\in\mathbb{R}} \left|\frac{y}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{y^2}{8}}\right| = C < +\infty;$$
$$|\mathbb{P}[\mathcal{N}(0,1)\geq\beta] - \mathbb{P}[\mathcal{N}(0,1)\geq y]| = \left|F_y(1+o((t_n)^{-1/6})) - F_y(1)\right| \leq \frac{C}{(t_n)^{1/6}} = o(1).$$

This ends the proof of Theorem 2.8, with the refinement that y can be as large as a $o((t_n)^{1/6})$. □

Remark 2.11. Theorems 2.4 and 2.8 generalize the usual central limit theorem and Bahadur-Rao's estimates (see *e.g.* [DZ98, Theorem 3.7.4]) for sums of lattice-valued independent and identically distributed random variables. Indeed, one can consider that a sum $S_n = X_1 + \cdots + X_n$ of i.i.d. variables "converges mod-X with parameters $t_n = n$ and limiting function $\psi(z) = 1$", since $\mathbb{E}[\mathrm{e}^{zS_n}] = (\mathbb{E}[\mathrm{e}^{zX}])^n$. Notice though that $\mathbb{E}[\mathrm{e}^{zX}]$ might vanish and does not correspond to an infinitely divisible law, whence the quotes in the previous statement. However, one can follow exactly the same proofs to recover as a particular case the Bahadur-Rao estimates. The same remark will apply in the non-lattice case; for this case we refer also to [Ney83, Ilt95] which give analogue results of "precise" large deviations (that is to say that one estimates directly $\mathbb{P}[X_n \in t_n B]$ instead of $\frac{1}{t_n}\log \mathbb{P}[X_n \in t_n B]$). In a sense, the content of this paper is the missing box in the following diagram:

$$\begin{tikzcd}
\begin{array}{c}\text{Cramér's large deviations}\\ \text{for sums of i.i.d. variables}\end{array} \arrow[r, "\text{precise estimates}"] \arrow[d, "\text{generalization}"'] & \text{Bahadur-Rao theorem} \arrow[d, "\text{generalization}"] \\
\begin{array}{c}\text{Ellis-Gärtner theorem}\\ \text{(conditions on } \frac{1}{t_n}\log\varphi_n(\cdot))\end{array} \arrow[r, "\text{precise estimates}"'] & \boxed{???}
\end{tikzcd}$$

Notice however that the regime (order of renormalization) in which the precise estimates of probabilities will be given can be different from the usual regime of large deviations. In other words, we shall sometimes obtain the so-called *moderate deviations* of the sequence of random variables. For instance, for sums of i.i.d. random variables, the central limit theorem gives estimates in the regime $O(\sqrt{n})$, Cramér's theorem gives non-precise (logarithmic) estimates in the regime $O(n)$, and we will state results in the regime $O(n^{2/3})$, or even $o(n^{3/4})$, see Theorem 3.2.

Example 2.12. Suppose that $(X_n)_{n\in\mathbb{N}}$ is mod-Poisson convergent, that is to say that $\eta(z) = \mathrm{e}^z - 1$. The expansion reads then as follows:

$$\mathbb{P}[X_n = t_n x] = \frac{\mathrm{e}^{t_n(x-1-x\log x)}}{\sqrt{2\pi x t_n}}\left(\psi(h) + \frac{\psi'(h) - 3\psi''(h) - \psi(h)}{6xt_n} + O\left(\frac{1}{(t_n)^2}\right)\right)$$

with $h = \log x$. For instance, if X_n is the number of cycles of a random permutation in $\mathfrak{S}_n$, then for $x > 0$,

$$\mathbb{P}[X_n = \lfloor x(\log n)\rfloor] = \frac{n^{-(x\log x - x + 1)}}{\sqrt{2\pi x\,\log n}}\,\frac{1}{\Gamma(x)}\big(1 + o(1)\big),$$

and for $x > 1$,

$$\mathbb{P}[X_n \geq x(\log n)] = \frac{n^{-(x\log x - x + 1)}}{\sqrt{2\pi x\,\log n}}\,\frac{x}{x-1}\,\frac{1}{\Gamma(x)}\big(1 + o(1)\big).$$

Example 2.13. Equation (5) is the probabilistic counterpart of the number-theoretic results of [Kub72, Rad09], see in particular Theorems 2.1 and 2.2 in [Rad09]. In §5.2, we shall explain how to recover the precise large deviation results of [Rad09] for arithmetic functions whose Dirichlet series can be studied with the Selberg-Delange method.

3. Large deviations in the non-lattice case

3.1. **Berry-Esseen estimates and large deviations in the scale $O(t_n)$.** In this section we prove the analogues of Theorems 2.4 and 2.8 when ϕ is not lattice-distributed; hence, $|\mathrm{e}^{\eta(\mathrm{i}u)}| < 1$ for any $u \neq 0$. In this setting, there is a formula equivalent to the one given in Lemma 2.3, namely,

$$\mathbb{P}[X > x] + \frac{1}{2}\,\mathbb{P}[X = x] = \lim_{R\to\infty}\left(\frac{1}{2\pi}\int_{-R}^{R}\frac{\mathrm{e}^{-x(h+\mathrm{i}u)}}{h+\mathrm{i}u}\,\varphi_X(h+\mathrm{i}u)\,du\right) \tag{6}$$

if $\varphi_X(h) = \mathbb{E}[\mathrm{e}^{hX}] < +\infty$ for $h > 0$. However, in order to manipulate this formula as in Section 2, one would need strong additional assumptions of integrability on the characteristic functions of the random variables X_n. On the other hand, for non-lattice distributions, one would like to know not only the precise asymptotics of probabilities $\mathbb{P}[X_n \geq t_n x]$, but more generally of probabilities

$$\mathbb{P}[X_n \in t_n B] \quad \text{with } B \text{ arbitrary Borelian subset of } \mathbb{R}.$$

To this purpose, we should combine classical arguments of the theory of large deviations (see [DZ98]), and the following lemma which replaces Equation (6) and is the main tool for estimates of probabilities $\mathbb{P}[X_n \geq t_n x]$.

Lemma 3.1 (Berry-Esseen expansion)**.** *Denote $F_n(x) = \mathbb{P}[X_n \leq t_n\eta'(0) + \sqrt{t_n\eta''(0)}\,x]$, and $p(y) = (2\pi)^{-1/2}\,\mathrm{e}^{-y^2/2}$ the density of a standard Gaussian variable. Under the assumptions of Definition 1.1, with ϕ non-lattice,*

$$F_n(x) = \left(\int_{-\infty}^{x}\left(1 + \frac{\psi'(0)}{\sqrt{t_n\eta''(0)}}\,y + \frac{\eta'''(0)}{6\sqrt{t_n(\eta''(0))^3}}\,(y^3 - 3y)\right)p(y)\,dy\right) + o\left(\frac{1}{\sqrt{t_n}}\right)$$

with the $o(\cdot)$ uniform on $\mathbb{R}$.

Proof. We use the same arguments as in the proof of [Fel71, Theorem XVI.4.1], but adapted to the assumptions of Definition 1.1. Given an integrable function f, its Fourier transform is $f^*(\zeta) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\zeta x}\, f(x)\, dx$. Consider a probability law $F(x) = \int_{-\infty}^{x} f(y)\, dy$ with vanishing expectation $(f^*)'(0) = 0$; and $G(x) = \int_{-\infty}^{x} g(y)\, dy$ a m-Lipschitz function with g^* continuously differentiable and

$$(g^*)'(0) = 0 \qquad ; \qquad \lim_{y\to-\infty} G(y) = 0 \qquad ; \qquad \lim_{y\to+\infty} G(y) = 1.$$

By [Fel71, Lemma XVI.3.2], for any $x \in \mathbb{R}$ and any $T > 0$,

$$|F(x) - G(x)| \leq \frac{1}{\pi} \int_{-T}^{T} \left| \frac{f^*(\zeta) - g^*(\zeta)}{\zeta} \right| d\zeta + \frac{24m}{\pi T}.$$

Notice that this is true even when F does not admit a density f (see also §4.2.3). We shall apply this result to the functions

$$F(x) = F_n(x) = \text{cumulative distribution function of } Y_n = \frac{X_n - t_n\eta'(0)}{\sqrt{t_n\eta''(0)}};$$

$$G(x) = G_n(x) = \int_{-\infty}^{x} \left(1 + \frac{\psi'(0)}{\sqrt{t_n\eta''(0)}}\, y + \frac{\eta'''(0)}{6\sqrt{t_n(\eta''(0))^3}}\, (y^3 - 3y) \right) p(y)\, dy.$$

The asymptotic behavior of $f_n^*(\zeta)$ is given by the local uniform convergence $\psi_n(z) \to \psi(z)$:

$$\begin{aligned}
\mathbb{E}\left[\mathrm{e}^{z\left(\frac{X_n - t_n\eta'(0)}{\sqrt{t_n\eta''(0)}}\right)}\right] &= \exp\left(-z\sqrt{\frac{t_n}{\eta''(0)}}\, \eta'(0)\right) \times \varphi_n\left(\frac{z}{\sqrt{t_n\eta''(0)}}\right) \\
&= \exp\left(t_n\left(\eta\left(\frac{z}{\sqrt{t_n\eta''(0)}}\right) - \eta'(0)\, \frac{z}{\sqrt{t_n\eta''(0)}}\right)\right) \times \psi_n\left(\frac{z}{\sqrt{t_n\eta''(0)}}\right) \\
&= \exp\left(t_n\left(\eta\left(\frac{z}{\sqrt{t_n\eta''(0)}}\right) - \eta'(0)\, \frac{z}{\sqrt{t_n\eta''(0)}}\right)\right) \times \left(1 + \frac{\psi_n'(0)\, z}{\sqrt{t_n\eta''(0)}} + o\left(\frac{z}{\sqrt{t_n}}\right)\right) \\
&= \exp\left(\frac{z^2}{2} + \frac{\eta'''(0)\, z^3}{6\sqrt{t_n(\eta''(0))^3}} + |z|^2\, o\left(\frac{z}{\sqrt{t_n}}\right)\right) \times \left(1 + \frac{\psi'(0)\, z}{\sqrt{t_n\eta''(0)}} + o\left(\frac{z}{\sqrt{t_n}}\right)\right) \\
&= \mathrm{e}^{\frac{z^2}{2}}\left(1 + \frac{\psi'(0)\, z}{\sqrt{t_n\eta''(0)}} + \frac{\eta'''(0)\, z^3}{6\sqrt{t_n(\eta''(0))^3}} + (1 + |z|^2)\, o\left(\frac{z}{\sqrt{t_n}}\right)\right).
\end{aligned} \tag{7}$$

Beware that in the previous expansions, the only $o(\cdot)$ that we manipulate is a

$$o\left(\frac{z}{\sqrt{t_n}}\right) = \frac{|z|}{\sqrt{t_n}}\, \varepsilon\left(\frac{z}{\sqrt{t_n}}\right) \quad \text{with } \lim_{t\to 0} \varepsilon(t) = 0.$$

In particular, z might still go to infinity in this situation. To make everything clear we will continue to use the notation $\varepsilon(t)$ in the following. Up to the $o(\cdot)$, the last expression evaluated at $z = \mathrm{i}\zeta$ is the Fourier transform of g_n. Fix $0 < \delta < \Delta$ and take $T = \Delta\sqrt{t_n}$.

By Feller's lemma,

$$\begin{aligned}|F_n(x)-G_n(x)| &\leq \frac{1}{\pi}\int_{-\Delta\sqrt{t_n}}^{\Delta\sqrt{t_n}}\left|\frac{f_n^*(\zeta)-g_n^*(\zeta)}{\zeta}\right|d\zeta+\frac{24m}{\Delta\pi\sqrt{t_n}}\\ &\leq \frac{1}{\pi\sqrt{t_n}}\int_{-\delta\sqrt{t_n}}^{\delta\sqrt{t_n}}\mathrm{e}^{-\frac{\zeta^2}{2}}\,(1+|\zeta|^2)\,\varepsilon\!\left(\frac{\zeta}{\sqrt{t_n}}\right)d\zeta+\frac{24m}{\Delta\pi\sqrt{t_n}}\\ &\quad+\frac{1}{\pi\delta\sqrt{t_n}}\int_{[-\Delta\sqrt{t_n},\Delta\sqrt{t_n}]\setminus[-\delta\sqrt{t_n},\delta\sqrt{t_n}]}|f_n^*(\zeta)-g_n^*(\zeta)|\,d\zeta.\end{aligned}$$

In the right-hand side, the first part is an $\frac{\varepsilon(\delta)}{\sqrt{t_n}}$, and the second part is smaller than $\frac{M}{\Delta\sqrt{t_n}}$ for some constant M. Finally, the last integral goes to zero faster than any power of t_n. Indeed, for $|\zeta|\in[\delta\sqrt{t_n},\Delta\sqrt{t_n}]$,

$$|f_n^*(\zeta)|=\left|\varphi_n\left(\frac{\mathrm{i}\zeta}{\sqrt{t_n\eta''(0)}}\right)\right|\leq K(\Delta)\left(\max_{\sqrt{\eta''(0)}|u|\in[\delta,\Delta]}|\exp(\eta(\mathrm{i}u))|\right)^{t_n}\leq K(\Delta)\,(q_\delta)^{t_n}$$

with $q_\delta<1$, and the constant $K(\Delta)$ determined by the uniform convergence of ψ_n to ψ and by the behavior of ψ on the complex segment $[-\mathrm{i}\Delta/\sqrt{\eta''(0)},\mathrm{i}\Delta/\sqrt{\eta''(0)}]$. The same argument works for $|g_n^*(\zeta)|$; in the following the symbols $\widetilde{\ }$ indicate constants that take care simultaneously of $f_n^*(\zeta)$ and $g_n^*(\zeta)$. Fix $\varepsilon>0$, then δ such that $\varepsilon(\delta)<\varepsilon$ and $M\delta<\varepsilon$. Take $\Delta=\frac{1}{\delta}$; we get

$$|F_n(x)-G_n(x)|\leq\frac{2\varepsilon}{\sqrt{t_n}}+\frac{1}{\pi\delta\sqrt{t_n}}\,\widetilde{K}(\delta^{-1})\,(\widetilde{q}_\delta)^{t_n}\leq\frac{3\varepsilon}{\sqrt{t_n}}$$

for t_n big enough. This ends the proof of the lemma. □

Theorem 3.2. *Suppose ϕ non-lattice. If x is in the range of $\eta'_{|(0,c)}$, then*

$$\mathbb{P}[X_n\geq t_nx]=\frac{\exp(-t_nF(x))}{h\sqrt{2\pi t_n\eta''(h)}}\,\psi(h)\,(1+o(1))$$

where as usual h is defined by the implicit equation $\eta'(h(x))=x$.

Remark 3.3. Our theorem should be compared with [Hwa96, Theorem 1], which studies another regime of large deviations in the mod-ϕ setting, namely, when h goes to zero (or equivalently, $x\to\eta'(0)$).

Remark 3.4. The main difference between Theorems 2.4 and 3.2 is the replacement of the factor $\psi(h)/(1-\mathrm{e}^{-h})$ by $\psi(h)/h$; the same happens with Bahadur-Rao's estimates when going from lattice distributions to non-lattice distributions.

Remark 3.5. In order to ease their uses and generalizations to a multidimensional setting, we have stated Theorems 3.1 and 3.2 with a single term in the expansions, but again there is a recipe to push these expansions further. Recall that the Hermite polynomials are defined by

$$H_r(x)=(-1)^r\,\mathrm{e}^{\frac{x^2}{2}}\,\frac{\partial^r}{\partial x^r}\left(\mathrm{e}^{-\frac{x^2}{2}}\right).$$

The first Hermite polynomials are $H_0(x)=1$, $H_1(x)=x$, $H_2(x)=x^2-1$, $H_3(x)=x^3-3x$, *etc.* An important property of Hermite polynomials is that

$$(p(x)\,H_r(x))^*(\zeta)=(\mathrm{i}\zeta)^r\,\mathrm{e}^{-\frac{\zeta^2}{2}}; \tag{8}$$

it can be seen from the generating series of Hermite polynomials $\sum_{r=0}^{\infty} H_r(x)\, \frac{t^r}{r!} = \mathrm{e}^{tx - x^2/2}$. On the other hand, with $z = \mathrm{i}\zeta$, the quantity (7) appearing in the proof of Lemma 3.1 corresponds to the first terms of

$$f_n(\zeta) = \sum_{k=0}^{\infty} \frac{\alpha_k(\zeta)}{(t_n)^{k/2}},$$

where f_n is the same function as in Theorem 2.4, with $h = 0$. Each coefficient $\alpha_k(\zeta)$ is a polynomial in $(\mathrm{i}\zeta)$, and we denote $A_k(x)$ the image of $\alpha_k(\zeta)$ by the linear map $(\mathrm{i}\zeta)^r \mapsto H_r(x)$. Then, by using Equation (8) and adapting the proof, one sees that a better approximation of $F_n(x)$ is

$$\left(\int_{-\infty}^{x} \left(\sum_{k=0}^{2v} \frac{A_k(y)}{(t_n)^{k/2}}\right) p(y)\, dy\right) + o\left(\frac{1}{(t_n)^v}\right)$$

if one assumes that the speed of convergence is a $o((t_n)^{-v})$. This allows in Theorem 3.2 to obtain an expansion of

$$\mathbb{P}[X_n \geq t_n x]\, \exp(t_n F(x))\, (h\, \sqrt{2\pi t_n \eta''(h)})$$

in powers of $1/\sqrt{t_n}$, though one also needs for these computations the precise asymptotics of all the integrals $\int_0^{\infty} y^k\, \mathrm{e}^{-(y+a)^2/2}\, dy$ with k fixed and a going to infinity.

Proof of Theorem 3.2. The main trick, well known in the theory of large deviations, is to make the exponential change of probability

$$\mathbb{Q}[dy] = \frac{\mathrm{e}^{hy}}{\varphi_{X_n}(h)}\, \mathbb{P}[X_n \in dy];$$

we denote $\widetilde{X}_n$ a random variable following this law. The Y_n's have for generating functions $\varphi_{\widetilde{X}_n}(z) = \varphi_{X_n}(z+h)/\varphi_{X_n}(h)$, so they converge mod-$\phi'$, where ϕ' is the infinitely divisible distribution with characteristic function $\mathrm{e}^{\eta(h+z)-\eta(h)}$. The limiting function in the mod-ϕ' convergence is also replaced by $\psi(z+h)/\psi(h)$, but otherwise all the assumptions of Lemma 3.1 are satisfied. So, the distribution function $F_n(u)$ of $\frac{\widetilde{X}_n - t_n\eta'(h)}{\sqrt{t_n\eta''(h)}}$ is

$$G_n(u) = \int_{-\infty}^{u} \left(1 + \frac{\psi'(h)}{\psi(h)\, \sqrt{t_n\eta''(h)}}\, y + \frac{\eta'''(h)}{\sqrt{t_n(\eta''(h))^3}}\, (y^3 - 3y)\right) p(y)\, dy$$

up to a uniform $o(1/\sqrt{t_n})$. Then,

$$\begin{aligned}\mathbb{P}[X_n \geq t_n x] &= \int_{y=t_n x}^{\infty} \varphi_{X_n}(h)\, \mathrm{e}^{-hy}\, \mathbb{Q}(dy) = \varphi_{X_n}(h) \int_{u=0}^{\infty} \mathrm{e}^{-h\left(t_n\eta'(h) + \sqrt{t_n\eta''(h)}\, u\right)}\, dF_n(u) \\ &= \psi_n(h)\, \mathrm{e}^{-t_n F(x)} \int_{u=0}^{\infty} \mathrm{e}^{-h\sqrt{t_n\eta''(h)}\, u}\, dF_n(u).\end{aligned}$$

To compute the integral I, we choose the primitive of $dF_n(u)$ that vanishes at $u = 0$ and we make an integration by parts:

$$\begin{aligned}
I &= \left[\mathrm{e}^{-h\sqrt{t_n\eta''(h)}\,u}\,(F_n(u) - F_n(0))\right]_{u=0}^{\infty} + h\sqrt{t_n\eta''(h)}\int_{u=0}^{\infty} \mathrm{e}^{-h\sqrt{t_n\eta''(h)}\,u}\,(F_n(u) - F_n(0))\,du \\
&= h\sqrt{t_n\eta''(h)}\int_{u=0}^{\infty} \mathrm{e}^{-h\sqrt{t_n\eta''(h)}\,u}\left(G_n(u) - G_n(0) + o\left(\frac{1}{\sqrt{t_n}}\right)\right)du \\
&\simeq h\sqrt{t_n\eta''(h)}\iint_{0\leq y\leq u} \mathrm{e}^{-h\sqrt{t_n\eta''(h)}\,u}\left(1 + \frac{\psi'(h)\,y}{\psi(h)\,\sqrt{t_n\eta''(h)}} + \frac{\eta'''(h)\,(y^3-3y)}{\sqrt{t_n(\eta''(h))^3}}\right)p(y)\,dy\,du \\
&\simeq \int_{y=0}^{\infty} \mathrm{e}^{-h\sqrt{t_n\eta''(h)}\,y}\left(1 + \frac{\psi'(h)}{\psi(h)\,\sqrt{t_n\eta''(h)}}\,y + \frac{\eta'''(h)}{\sqrt{t_n(\eta''(h))^3}}\,(y^3-3y)\right)p(y)\,dy,
\end{aligned}$$

where on the two last lines the symbol $\simeq$ means that the remainder is a $o((t_n)^{-1/2})$. Also by integration by parts, one can compute recursively the estimates

$$\int_0^{\infty} \mathrm{e}^{-\frac{(y+a)^2}{2}}\,dy = \frac{\mathrm{e}^{-\frac{a^2}{2}}}{a}\left(1 - \frac{1}{a^2} + O\left(\frac{1}{a^4}\right)\right) \quad ; \quad \int_0^{\infty} y^2\,\mathrm{e}^{-\frac{(y+a)^2}{2}}\,dy = O\left(\frac{\mathrm{e}^{-\frac{a^2}{2}}}{a^3}\right)$$

$$\int_0^{\infty} y\,\mathrm{e}^{-\frac{(y+a)^2}{2}}\,dy = \frac{\mathrm{e}^{-\frac{a^2}{2}}}{a^2}\left(1 + O\left(\frac{1}{a^2}\right)\right) \quad ; \quad \int_0^{\infty} y^3\,\mathrm{e}^{-\frac{(y+a)^2}{2}}\,dy = O\left(\frac{\mathrm{e}^{-\frac{a^2}{2}}}{a^2}\right)$$

for a going to infinity, and this shows that the only contribution in the integral that is not a $o((t_n)^{-1/2})$ is

$$\int_{y=0}^{\infty} \mathrm{e}^{-h\sqrt{t_n\eta''(h)}\,y}\,p(y)\,dy = \frac{1}{h\sqrt{2\pi t_n\eta''(h)}} + o\left(\frac{1}{\sqrt{t_n}}\right).$$

This ends the proof since $\psi_n(h) \to \psi(h)$ locally uniformly. □

3.2. **Precise moderate deviations for random variables with control on cumulants.** Let us come back to the example of centered random variables S_n having cumulants $\kappa^{(r)}(S_n)$ bounded in absolute value by $(Cr)^r\,\alpha_n\,(\beta_n)^r$, where $(\alpha_n)_{n\in\mathbb{N}}$ and $(\beta_n)_{n\in\mathbb{N}}$ are two arbitrary sequences, α_n going to infinity — in the example of the introduction 1.7, we had $\alpha_n = n^a$ and $\beta_n = n^b$, which is the most common case. We also suppose that $\kappa^{(2)}(S_n) = \sigma^2\,\alpha_n\,(\beta_n)^2\,(1 + o((\alpha_n)^{-5/12}))$ and that $\kappa^{(3)}(S_n) = L\,\alpha_n\,(\beta_n)^3\,(1 + o((\alpha_n)^{-1/6}))$. Then, one observes the following asymptotic regimes:

(1) For any sequence $s_n = o(1)$,

$$\mathbb{P}\left[S_n \geq s_n\,\beta_n\,(\alpha_n)^{\frac{1}{2}}\right] = \frac{1}{2}\big(1 + o(1)\big).$$

(2) More generally, for any sequence $s_n = O(1)$,

$$\mathbb{P}\left[S_n \geq s_n\,\beta_n\,(\alpha_n)^{\frac{1}{2}}\right] = \mathbb{P}[\mathcal{N}(0,\sigma^2) \geq s_n]\big(1 + o(1)\big).$$

This follows immediately from Lemma 3.1, and one even knows that the remainder is in fact a $O((\alpha_n)^{-1/6})$.

(3) Actually, the previous expansion holds for $s_n = o((\alpha_n)^{1/6})$, see [Hwa96], and our Proposition 2.8 for the analog in the lattice case. So, for any such sequence going to infinity,

$$\mathbb{P}\left[S_n \geq s_n\,\beta_n\,(\alpha_n)^{\frac{1}{2}}\right] = \frac{\exp\left(-\frac{(s_n)^2}{2\sigma^2}\right)}{s_n\sqrt{2\pi\sigma^2}}\big(1+o(1)\big).$$

(4) At the next level, the function $\psi(x) = \exp(\frac{Lx^3}{6})$ comes into play, so for any sequence $s_n = O(1)$ bounded from both sides by positive constants,

$$\mathbb{P}\left[S_n \geq s_n\,\beta_n\,(\alpha_n)^{\frac{2}{3}}\right] = \frac{\exp\left(-(\alpha_n)^{\frac{1}{3}}\frac{(s_n)^2}{2\sigma^2}\right)}{s_n\,(\alpha_n)^{\frac{1}{6}}\,\sqrt{2\pi\sigma^2}}\,\psi(s_n)\big(1+o(1)\big).$$

Indeed, we can apply Theorem 3.2 to the sequence $S_n/(\beta_n\,(\alpha_n)^{1/3})$, which converges mod-Gaussian with speed $(\alpha_n)^{1/3}\,\sigma^2$ and limiting function ψ.

One may then ask up to which order the previous expansion can be pushed. Up to a renormalization of the random variables, one can suppose $\sigma^2 = 1$. Set then $X_n = S_n/(\beta_n\,(\alpha_n)^{1/3})$. Notice that as long as $z_n = o((\alpha_n)^{1/12})$, the estimate

$$\varphi_{X_n}(z_n) = \exp\left((\alpha_n)^{\frac{1}{3}}\frac{(z_n)^2}{2} + \frac{L(z_n)^3}{6}\right)\big(1+o(1)\big)$$

holds, because the terms of order $r \geq 4$ in the series expansion of $\log\varphi_{X_n}$ all go to zero — this is the same computation as after Definition 1.1, but noticing that the $O((\alpha_n)^{-1/3})$ in the estimate of the remainder is in fact a $O((z_n)^4\,(\alpha_n)^{-1/3})$, that is, a $o(1)$ under the previous assumptions. This leads us to try to push the expansion up to $x = x_n\,(\alpha_n)^{1/12}$, with $x_n = o(1)$. In the following we also suppose $x \to \infty$ since the other situation is already known. If one makes the change of probability $\mathbb{P}[Y_n \in dy] = \frac{\mathrm{e}^{xy}}{\varphi_{X_n}(x)}\,\mathbb{P}[X_n \in dy]$, then the generating function of Y_n writes

$$\begin{aligned}
\varphi_{Y_n}(z) &= \frac{\varphi_{X_n}(x+z)}{\varphi_{X_n}(x)} = \exp\left(\sum_{r=2}^{\infty}\frac{\kappa^{(r)}(X_n)}{r!}\,((x+z)^r - x^r)\right)\\
\log\varphi_{Y_n}(z) &= \frac{(\alpha_n)^{\frac{1}{3}}}{2}\left(z^2 + 2z\,x_n\,(\alpha_n)^{\frac{1}{12}}\right)\left(1+o\Big((\alpha_n)^{-\frac{5}{12}}\Big)\right)\\
&\quad+\frac{L}{6}\left(z^3 + 3z^2\,x_n\,(\alpha_n)^{\frac{1}{12}} + 3z\,(x_n)^2\,(\alpha_n)^{\frac{1}{6}}\right)\left(1+o\Big((\alpha_n)^{-\frac{1}{6}}\Big)\right)\\
&\quad+\sum_{r=4}^{\infty}\frac{\kappa^{(r)}(X_n)}{r!}\,((x+z)^r - x^r). \qquad (9)
\end{aligned}$$

According to the previous discussion, the sum of terms with $r \geq 4$ goes to zero. So, locally uniformly,

$$\log\varphi_{Y_n}(z) = \left((\alpha_n)^{\frac{5}{12}}\,x_n + \frac{L\,(\alpha_n)^{\frac{1}{6}}\,(x_n)^2}{2}\right)z + \left(\frac{(\alpha_n)^{\frac{1}{3}} + L\,(\alpha_n)^{\frac{1}{12}}\,x_n}{2}\right)z^2 + \frac{L}{6}\,z^3 + o(1),$$

and if $Z_n = Y_n - (\alpha_n)^{\frac{5}{12}}\,x_n - \frac{L\,(\alpha_n)^{\frac{1}{6}}\,(x_n)^2}{2}$, then the Z_n's converge in the mod-Gaussian sense with parameter $(\alpha_n)^{\frac{1}{3}} + L\,(\alpha_n)^{\frac{1}{12}}\,x_n$ and limiting function $\exp(\frac{L\,z^3}{6})$. One can therefore

apply Lemma 3.1, and

$$\mathbb{P}\Big[S_n \geq x_n\,\beta_n\,(\alpha_n)^{\frac{3}{4}}\Big] = \mathbb{P}\Big[X_n \geq (\alpha_n)^{\frac{1}{3}}\,x\Big] = \varphi_{X_n}(x)\int_{y=(\alpha_n)^{1/3}\,x}^{\infty} \mathrm{e}^{-xy}\,\mathbb{P}[Y_n \in dy]$$

$$= \varphi_{X_n}(x)\,\mathrm{e}^{-\left((\alpha_n)^{1/2}(x_n)^2+\frac{L\,(\alpha_n)^{1/4}(x_n)^3}{2}\right)}\int_{u=-\frac{L\,(\alpha_n)^{1/6}\,(x_n)^2}{2}}^{\infty} \mathrm{e}^{-xu}\,\mathbb{P}[Z_n \in du]$$

$$= \psi_n(x)\,\mathrm{e}^{-(\alpha_n)^{1/3}\frac{x^2}{2}}\,R_n,$$

where R_n is given by

$$R_n = \mathrm{e}^{-\frac{L\,(\alpha_n)^{1/4}(x_n)^3}{2}}\int_{w=-\frac{L\,(x_n)^2}{2\sqrt{1+L\,(\alpha_n)^{-1/12}\,x_n}}}^{\infty} \mathrm{e}^{-\left(x_n\,(\alpha_n)^{1/4}\,\sqrt{1+L\,(\alpha_n)^{-1/12}\,x_n}\right)w}\,F_n(dw).$$

This integral R_n is now computed exactly as before, and we obtain in the end the following result, to be compared to Edgeworth's expansion in the central limit theorem, see for instance [Fel71, XVI.2 Theorem 2 and XVI.4 Theorem 3], and also the results of [DE12].

Corollary 3.6. *Let $(S_n)_{n\in\mathbb{N}}$ be a sequence of centered real-valued random variables such that $|\kappa^{(r)}(S_n)| \leq (Cr)^r\,\alpha_n\,(\beta_n)^r$ with $\alpha_n \to \infty$ and*

$$\kappa^{(2)}(S_n) = \sigma^2\,\alpha_n\,(\beta_n)^2\,(1+o((\alpha_n)^{-5/12})) \quad ; \quad \kappa^{(3)}(S_n) = L\,\alpha_n\,(\beta_n)^3\,(1+o((\alpha_n)^{-1/6})).$$

If $(\alpha_n)^{1/2} \ll T \ll (\alpha_n)^{3/4}$, then

$$\mathbb{P}\left[\frac{S_n}{\beta_n} \geq T\right] = \frac{\mathrm{e}^{-\frac{T^2}{2\sigma^2\,\alpha_n}}}{\sqrt{2\pi\sigma^2\frac{T^2}{\alpha_n}}}\,\exp\!\left(\frac{L\,T^3}{6\,\sigma^3\,(\alpha_n)^2}\right)\big(1+o(1)\big),$$

the second term in the expansion being meaningful for $(\alpha_n)^{2/3} \lesssim T \ll (\alpha_n)^{3/4}$.

Corollary 3.6 hints at a possible expansion up to any order $T = o((\alpha_n)^{1-\varepsilon})$, and indeed, it is a particular case of the results given by Rudzkis, Saulis and Statulevičius in [RSS78, SS91], see in particular [SS91, Lemma 2.3]. Suppose that

$$|\kappa^{(r)}(S_n)| \leq (Cr)^r\,\alpha_n\,(\beta_n)^r \quad ; \quad \kappa^{(r)}(S_n) = K(r)\,\alpha_n\,(\beta_n)^r\,(1+O((\alpha_n)^{-1}))$$

the second estimate holding for any $r \leq p$; we denote $\sigma^2 = K(2)$. In this setting, one can push the expansion up to order $o((\alpha_n)^{1-1/p})$. Indeed, define recursively for a sequence of cumulants $(\kappa^{(r)})_{r\geq 2}$ the coefficients of the Petrov-Cramér series $\lambda^{(r)} = -b_{r-1}/r$, with

$$\sum_{r=1}^{j}\frac{\kappa^{(r+1)}}{r!}\left(\sum_{\substack{j_1+\cdots+j_r=j\\ j_i\geq 1}} b_{j_1}b_{j_2}\cdots b_{j_r}\right) = \delta_{j,1}.$$

For instance, $\lambda^{(2)} = -\frac{1}{2}$, $\lambda^{(3)} = \frac{\kappa^{(3)}}{6}$, $\lambda^{(4)} = \frac{\kappa^{(4)}-3(\kappa^{(3)})^2}{24}$, *etc.* The appearance of these coefficients can be guessed by trying to push the previous technique to higher order; in particular, the simple form of $\lambda^{(3)}$ is related to the fact that the only term in z^3 in the expansion (9) is $\frac{\kappa^{(3)}}{6}$. If for the $\kappa^{(r)}$'s one has estimates of order $(\alpha_n)^{1-r/2}(1+O((\alpha_n)^{-1}))$, then one has the same estimates for the $\lambda^{(r)}$'s, so there exists coefficients $L(r)$ such that

$$\lambda^{(r)}\left(\frac{S_n}{\beta_n\,(\alpha_n)^{\frac{1}{2}}}\right) = L(r)\,(\alpha_n)^{1-r/2}\,(1+O((\alpha_n)^{-1})).$$

Take then $T = x\,(\alpha_n)^{\frac{p-1}{p}}$ with $x = O(1)$, but $T\,(\alpha_n)^{-1/2}$ going to infinity; Lemma 2.3 of [SS91] ensures that

$$\begin{aligned}\mathbb{P}\left[\frac{S_n}{\beta_n} \geq T\right] &= \frac{\mathrm{e}^{-\frac{T^2}{2\sigma^2\,\alpha_n}}}{\sqrt{2\pi\sigma^2 \frac{T^2}{\alpha_n}}}\,\exp\left(\sum_{r=3}^{p} \lambda^{(r)}\left(\frac{T}{\sigma\,(\alpha_n)^{1/2}}\right)^r\right)\big(1+o(1)\big)\\ &= \frac{\mathrm{e}^{-\frac{T^2}{2\sigma^2\,\alpha_n}}}{\sqrt{2\pi\sigma^2 \frac{T^2}{\alpha_n}}}\,\exp\left(\sum_{r=3}^{p} \frac{L(r)\,T^r}{\sigma^r\,(\alpha_n)^{r-1}}\right)\big(1+o(1)\big)\end{aligned}$$

Unfortunately, in this estimate, one cannot go up to $T = O(\alpha_n)$, because the generating series of the $L(r)$'s does not make sense. Indeed, unless one starts already with Gaussian variables (that is to say that $K(r) = 0$ for all $r \geq 3$), this series has radius of convergence zero, since otherwise, by the classical (non-precise) large deviations, one would have necessarily $\sum_{r\geq 3}^{\infty} L(r)\left(\frac{z}{\sigma}\right)^r = 0$, whence a contradiction. Nevertheless, the discussion shows that the method of cumulants of Rudzkis, Saulis and Statulevičius can be thought of as a particular case (and refinement in this setting) of the notion of mod-ϕ convergence. Also, it shows that one cannot expect precise large deviations in the regime $T = O(\alpha_n)$ (or even $o(\alpha_n)$) by a cumulant method without an adequate "renormalization theory" of the Cramér-Petrov series.

These estimates will be used extensively in the last sections of the paper. To conclude this paragraph, let us mention that in the same setting, one can state results similar to the law of the iterated logarithm. Let $(S_n)_{n\in\mathbb{N}}$ be a sequence of random variables that satisfy the hypotheses of Corollary 3.6 with $\alpha_n = n$ and $\beta_n = 1$; the second condition can be ensured by looking at $\frac{S_n}{\beta_n}$ instead of S_n, and the first condition is quite common in the setting of sums of random variables whose graph of dependence is of bounded degree, see the examples starting from Section 7. We suppose that the S_n's are defined on the same probability space, and we look at sequences γ_n such that almost surely,

$$\limsup_{n\to\infty} \frac{S_n}{\gamma_n} \leq 1.$$

The main difference between the following Proposition and the usual law of the iterated logarithm is that one does not make any assumption of independence. Such assumptions are common in this setting, or at least some conditional independence (for instance, a law of the iterated logarithm can be stated for martingales); see the survey [Bin86] or [Pet75, Chapter X]. We obtain a less precise result, but which does not depend at all on the way one realizes the random variables S_n. In other words, for every possible coupling of the S_n's, the following is true:

Proposition 3.7. *Under the previous assumptions,*

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2\sigma^2\,n\log n}} \leq 1 \quad \textit{almost surely.}$$

Proof. Notice the term $\log n$ instead of $\log\log n$ for the usual law of iterated logarithm. One uses of course Borel-Cantelli lemma, and compute

$$\mathbb{P}[S_n \geq \sqrt{2(1+\varepsilon)\sigma^2\,n\,\log n}] \simeq \frac{\mathrm{e}^{-(1+\varepsilon)\log n}}{\sqrt{4\pi\sigma^4(1+\varepsilon)\log n}} \leq \frac{1}{n^{1+\varepsilon}} \quad \text{for } n \text{ big enough.}$$

For any $\varepsilon > 0$, this is summable, so eventually $S_n < \sqrt{2(1+\varepsilon)\sigma^2\, n\, \log n}$. Since this is true for every ε,

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2\sigma^2\, n \log n}} \leq 1. \qquad \square$$

3.3. **A refinement of the Ellis-Gärtner theorem.** Theorem 3.2 is also the first step of a refinement of the Ellis-Gärtner theorem (see *e.g.* [DZ98, Theorem 2.3.6]) with precise estimates of the probabilities $\mathbb{P}[X_n \in t_n B]$. Until the end of this section, we make the following assumptions:

(1) The random variables X_n satisfy the hypotheses of Definition 1.1 with $c = +\infty$ (in particular, ψ is entire on $\mathbb{C}$).

(2) The Legendre-Fenchel transform F is essentially smooth, that is to say that it takes finite values on a non-empty closed interval I_F and that $\lim F'(x) = \lim h(x) = \pm\infty$ when x goes to a bound of the interval I_F (*cf.* [DZ98, Definition 2.3.5]).

The later point is verified if ϕ is a Gaussian or Poisson law, which are the most important examples.

Lemma 3.8. *Let C be a closed subset of $\mathbb{R}$. Either $\inf_{u\in C} F(u) = +\infty$, or $\inf_{u\in C} F(u) = m$ is attained and $\{x \in C \mid F(x) = \min_{u\in C} F(u)\}$ consists of one or two real numbers $a \leq b$, with $a < \eta'(0) < b$ if $a \neq b$.*

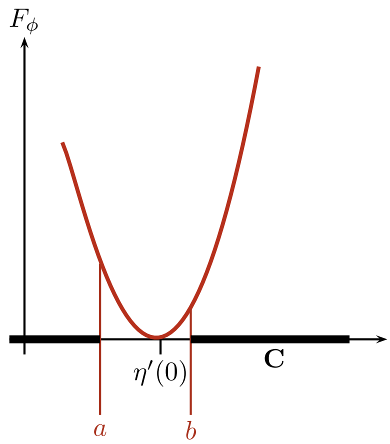


FIGURE 2. The infimum of F on an admissible closed set C is attained either at $a = \sup(C \cap (-\infty, \eta'(0)])$, or at $b = \inf(C \cap [\eta'(0), +\infty))$, or at both if $F(a) = F(b)$.

Proof. Recall that F is strictly convex, since its second derivative is $1/\eta''(h)$, which is the inverse of the variance of a non-constant random variable. Also, $\eta'(0)$ is the point where F attains its global minimum, and it is the expectation of the law ϕ. If $C \cap I_F = \emptyset$, then $F_{|C} = +\infty$ and we are in the first situation. Otherwise, $F_{|C}$ is finite at some points, so there exists $M \in \mathbb{R}_+$ such that $C \cap \{x \in \mathbb{R} \mid F(x) \leq M\} \neq \emptyset$. However, the set $\{x \in \mathbb{R} \mid F(x) \leq M\}$ is compact by the hypothesis of essential smoothness: it is closed as the reciprocal image of an interval $]-\infty, M]$ by a lower semi-continuous function, and bounded since $\lim_{x\to(I_F)^c} |F'(x)| = +\infty$. So, $C \cap \{x \in \mathbb{R} \mid F(x) \leq M\}$ is a non-empty

compact set, and the lower semi-continuous F attains its minimum on it, which is also $\min_{u\in C} F(u)$. Then, if $a \leq b$ are two points in C such that $F(a) = F(b) = \min_{u\in C} F(u)$, then by strict convexity of F, $F(x) < F(a)$ for all $x \in (a,b)$, hence, $(a,b) \subset C^{\mathrm{c}}$. Also, $F(x) > F(a)$ if $x \notin [a,b]$, so either $a = b$, or $\eta'(0) \in (a,b)$. □

We take the usual notations B^{o} and $\overline{B}$ for the interior and the closure of a subset $B \subset \mathbb{R}$. Call *admissible* a (Borelian) subset $B \subset \mathbb{R}$ such that $F_{|B} \neq +\infty$, and denote then $F(B) = \inf_{u\in B} F(u) = \min_{u\in\overline{B}} F(u)$, and $B_{\min} = \{a \in \overline{B} \mid F(a) = F(B)\}$; according to the previous discussion, $B_{\min}$ consists of one or two elements.

Theorem 3.9. *Let B be a Borelian subset of $\mathbb{R}$.*

(1) If B is admissible, then

$$\limsup_{n\to\infty} \left(\sqrt{2\pi t_n}\, \exp(t_n F(B))\, \mathbb{P}[X_n \in t_n B]\right) \leq \begin{cases} \sum_{a\in B_{\min}} \frac{\psi(h(a))}{(1-\mathrm{e}^{-|h(a)|})\sqrt{\eta''(h(a))}} \\ \sum_{a\in B_{\min}} \frac{\psi(h(a))}{|h(a)|\sqrt{\eta''(h(a))}} \end{cases}$$

the distinction of cases corresponding to ϕ lattice or non-lattice distributed. The sum on the right-hand side consists in one or two terms — it is considered infinite if $a = \eta'(0) \in B_{\min}$.

(2) If B is not admissible, then for any positive real number M,

$$\lim_{n\to\infty} \left(\exp(t_n M)\, \mathbb{P}[X_n \in t_n B]\right) = 0.$$

Proof. For the second part, one knows that $\varphi_n(x)\exp(-t_n\eta(x))$ converges to $\psi(x)$ which does not vanish on the real line, so by taking the logarithms,

$$\lim_{n\to\infty} \frac{\log \varphi_n(x)}{t_n} = \eta(x).$$

Then, Ellis-Gärtner theorem holds since F is supposed essentially smooth. So,

$$\limsup_{n\to\infty} \frac{\log \mathbb{P}[X_n \in t_n B]}{t_n} \leq -F(B),$$

and if B is not admissible, then the right-hand side is $-\infty$ and (2) follows immediately.

For the first part, suppose for instance ϕ non-lattice distributed. Take C a closed admissible subset, and assume $\eta'(0) \notin C$ — otherwise the upper bound in (1) is $+\infty$ and the inequality is trivially satisfied. Since C^{c} is an open set, there is an open interval $(a,b) \subset C^{\mathrm{c}}$ containing $\eta'(0)$, and which we can suppose maximal. Then a and b are in C as soon as they are finite, and $C \subset (-\infty,a] \sqcup [b,+\infty)$. Moreover, by strict convexity of F, the minimal value $F(C)$ is necessarily attained at a or b. Suppose for instance $F(a) = F(b) = F(C)$ — the other situations are entirely similar. Then,

$$\begin{aligned}\mathbb{P}[X_n \in t_n C] &\leq \mathbb{P}[X_n \leq t_n a] + \mathbb{P}[X_n \geq t_n b] \\ &\leq \exp(-t_n F(C)) \left(\frac{\psi(h(a))}{-h(a)\sqrt{2\pi t_n \eta''(h(a))}} + \frac{\psi(h(b))}{h(b)\sqrt{2\pi t_n \eta''(h(b))}}\right)(1+o(1))\end{aligned}$$

by using Theorem 3.2 for $\mathbb{P}[X_n \geq t_n b]$, and also for $\mathbb{P}[X_n \leq t_n a] = \mathbb{P}[-X_n \geq -t_n a]$ — the random variables $-X_n$ satisfy the same hypotheses as the X_n's with $\eta(x)$ replaced by $\eta(-x)$, $\psi(x)$ replaced by $\psi(-x)$, *etc.* This proves the upper bound when B is closed, and

since $F(B) = F(\overline{B})$ by lower semi-continuity of B and $B_{\min} = (\overline{B})_{\min}$, the result extends immediately to arbitrary admissible Borelian subsets. $\square$

One can then ask for an asymptotic lower bound on $\mathbb{P}[X_n \in t_n B]$, and in view of the classical theory of large deviations, this lower bound should be related to open sets and to the exponent $F(B^{\mathrm{o}})$. Unfortunately, the result takes a less interesting form than Theorem 3.9. If B is a Borelian subset of $\mathbb{R}$, denote B^δ the union of the open intervals $(x, x+\kappa)$ of width $\kappa \geq \delta$ that are included into B. The interior $O = B^{\mathrm{o}}$ is a disjoint union of a countable collection of open intervals, and also the increasing union $\bigcup_{\delta>0} B^\delta$.

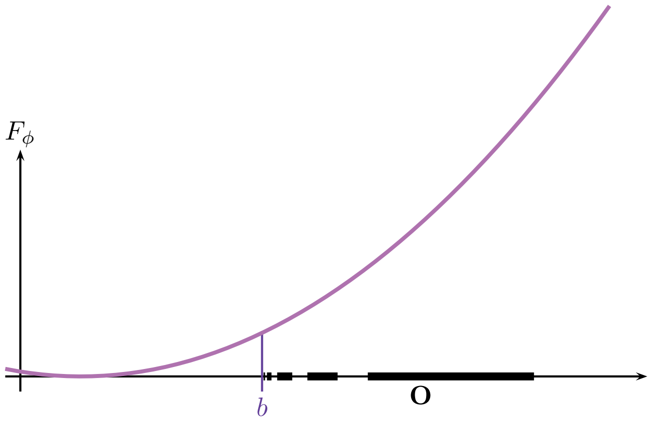


FIGURE 3. In some problematic situations, one is only able to prove a non-precise lower bound for large deviations.

However, the topology of B^{o} may be quite intricate in comparison to the one of the B^δ's, as some points can be points of accumulation of open intervals included in B and of width going to zero (see Figure 3). This phenomenon prevents us to state a precise lower bound when one of this point of accumulation is $a = \sup(B^{\mathrm{o}} \cap (-\infty, \eta'(0)])$ or $b = \inf(B^{\mathrm{o}} \cap [\eta'(0), +\infty))$. Nonetheless, the following is true:

Theorem 3.10. *For an admissible Borelian set B,*

$$\liminf_{\delta\to 0} \liminf_{n\to\infty} \left(\sqrt{2\pi t_n}\, \exp(t_n F(B^\delta))\, \mathbb{P}[X_n \in t_n B]\right) \geq \begin{cases} \sum_{a\in(B^{\mathrm{o}})_{\min}} \frac{\psi(h(a))}{(1-\mathrm{e}^{-|h(a)|})\sqrt{\eta''(h(a))}} \\ \sum_{a\in(B^{\mathrm{o}})_{\min}} \frac{\psi(h(a))}{|h(a)|\sqrt{\eta''(h(a))}}, \end{cases}$$

with again the distinction of cases lattice/non-lattice. In particular, the right-hand side in Theorem 3.9 is the limit of $\sqrt{2\pi t_n}\, \exp(t_n F(B))\, \mathbb{P}[X_n \in t_n B]$ as soon as $F(B^\delta) = F(B)$ for some $\delta > 0$.

Proof. Again we deal with the non-lattice case, and we suppose for instance that $(B^{\mathrm{o}})_{\min}$ consists of one point $b = \inf(B^{\mathrm{o}} \cap [\eta'(0), +\infty))$, the other situations being entirely similar. As δ goes to 0, B^δ increases towards $B^{\mathrm{o}} = \bigcup_{\delta>0} B^\delta$, so the infimum $F(B^\delta)$ decreases and the quantity

$$L(\delta) = \liminf_{n\to\infty} \left(\sqrt{2\pi t_n}\, \exp(t_n F(B^\delta))\, \mathbb{P}[X_n \in t_n B]\right)$$

is decreasing in δ. Actually, if $b^\delta = \inf(B^\delta \cap [\eta'(0), +\infty))$, then for δ small enough $F(B^\delta) = F(b^\delta)$, so $\lim_{\delta\to 0} F(B^\delta) = F(B^{\mathrm{o}})$ by continuity of F. On the other hand,

$$R(\delta) = \frac{\psi(h(b^\delta))}{h(b^\delta)\sqrt{\eta''(h(b^\delta))}}$$

goes to the same quantity with b instead of b^δ. Hence, it suffices to show that for δ small enough, $L(\delta) \geq R(\delta)$. However, by definition of B^δ, the open interval $(b^\delta, b^\delta + \delta)$ is included into B, so

$$\begin{aligned}\mathbb{P}[X_n \in t_n B] &\geq \mathbb{P}[X_n \in t_n B^\delta] \geq \mathbb{P}[X_n > t_n b^\delta] - \mathbb{P}[X_n \geq t_n(b^\delta + \delta)] \\ &\geq \frac{\psi(h(b^\delta))\,\mathrm{e}^{-t_n F(b^\delta)}}{h(b^\delta)\sqrt{2\pi t_n \eta''(h(b^\delta))}}\big(1+o(1)\big) - \frac{\psi(h(b^\delta+\delta))\,\mathrm{e}^{-t_n F(b^\delta+\delta)}}{h(b^\delta+\delta)\sqrt{2\pi t_n \eta''(h(b^\delta+\delta))}}\big(1+o(1)\big) \\ &\geq \frac{\psi(h(b^\delta))\,\mathrm{e}^{-t_n F(b^\delta)}}{h(b^\delta)\sqrt{2\pi t_n \eta''(h(b^\delta))}}\big(1+o(1)\big)\end{aligned}$$

since the second term on the second line is negligible in comparison to the first term — $F(b^\delta + \delta) > F(b^\delta)$. This ends the proof. ☐

4. Multi-dimensional extensions

In the following, vectors in $\mathbb{R}^d$ are denoted by bold letters $\mathbf{V}, \mathbf{W}, \ldots$, and their coordinates are indexed by exponents, so $\mathbf{V} = (\mathbf{V}^{(1)}, \ldots, \mathbf{V}^{(d)})$. This convention enables us to manipulate sequences $(\mathbf{V}_n)_{n\in\mathbb{N}}$ of (random) vectors in $\mathbb{R}^d$ without ambiguity.

4.1. **Sum of a Gaussian noise and an independent random variable.**

The multi-dimensional extension of the previous results is geometrically non-trivial, and we shall restrict ourselves to the following situation, which is the mod-Gaussian setting:

(a) A sequence of $\mathbb{R}^d$-valued random variables $(\mathbf{X}_n)_{n\in\mathbb{N}}$ is said to converge in the mod-Gaussian sense with parameters $A_n \in \mathrm{GL}(\mathbb{R}^d)$ and limiting function ψ if locally uniformly on $\mathbb{C}^d$,

$$\exp\left(-\frac{1}{2}\sum_{i=1}^{d}((A_n^*\mathbf{z})^{(i)})^2\right)\,\mathbb{E}\left[\mathrm{e}^{\sum_{i=1}^d \mathbf{X}_n^{(i)}\mathbf{z}^{(i)}}\right] \to \psi(\mathbf{z}),$$

with ψ analytic in d variables and non-vanishing on its real domain, and $\Sigma_n = (A_n)^{-1}$ going to zero.

(b) By replacing A_n by $\sqrt{A_n A_n^*}$, it does not cost anything in our definition to suppose that A_n is symmetric definite positive.

(c) Our goal is then to find the precise asymptotics of the probabilities $\mathbb{P}[\mathbf{X}_n \in (A_n)^2 B]$, where B is a Borelian subset of $\mathbb{R}^d$. In the one-dimensional case, $A_n = \sqrt{t_n}$ and we can rewrite the estimate of Theorem 3.9 as

$$\limsup_{n\to\infty}\left(\sqrt{\frac{2\pi}{t_n}}\,(t_n b)\,\exp\left(\frac{t_n b^2}{2}\right)\,\mathbb{P}[X_n \in t_n B]\right) \leq \int_{B_{\min}} \psi(x)\,N(dx)$$

where $b = \inf\{|x|\ ,\ x \in B\}$, and N is the counting measure on $B_{\min}$, which consists of one or two points. Moreover this estimate is sharp as soon as B contains small open intervals $[b, b+\delta)$ or $(-b-\delta, -b]$ (according to the type of $B_{\min}$).

Example 4.1. In order to make a correct conjecture, let us consider a toy model. We suppose here that $\mathbf{X}_n = \mathbf{Y} + A_n\mathbf{G}$, where $\mathbf{G}$ is a d-dimensional standard Gaussian variable,

the A_n's are symmetric and positive definite, and $\mathbf{Y}$ is a fixed random variable independent of the Gaussian $\mathbf{G}$ and with generating series $\psi(\mathbf{z})$. A simple computation yields

$$\mathbb{E}\left[\mathrm{e}^{\sum_{i=1}^d \mathbf{X}_n^{(i)}\mathbf{z}^{(i)}}\right] = \mathbb{E}[\mathrm{e}^{\langle A_n\mathbf{G}\,|\,\mathbf{z}\rangle}]\,\psi(\mathbf{z}) = \mathbb{E}[\mathrm{e}^{\langle \mathbf{G}\,|\,A_n^*\mathbf{z}\rangle}]\,\psi(\mathbf{z}) = \exp\left(\frac{1}{2}\sum_{i=1}^d((A_n^*\mathbf{z})^{(i)})^2\right)\psi(\mathbf{z}),$$

so we are typically in the mod-Gaussian setting. We assume in the following that:

(d) The random variable $\mathbf{Y}$ has a density $\nu(\mathbf{y})$ with respect to the Lebesgue measure, and is a bounded random variable. In particular, $\psi(\mathbf{b}) \leq \mathrm{e}^{C\|A\mathbf{b}\|}$ for some constant C.

The probabilities that we want to estimate write then as:

$$\begin{aligned}
\mathbb{P}[\mathbf{X}_n \in (A_n)^2 B] &= \frac{1}{(2\pi)^{d/2}}\int_{\mathbf{x}\in\mathbb{R}^d}\int_{\mathbf{y}\in\mathbb{R}^d} \mathrm{e}^{-\frac{\mathbf{x}^*\mathbf{x}}{2}}\,\mathbb{1}_{(\mathbf{y}+A_n\mathbf{x}\,\in\,(A_n)^2B)}\,\nu(\mathbf{y})\,d\mathbf{x}\,d\mathbf{y}\\
&= \frac{\det A_n}{(2\pi)^{d/2}}\int_{\mathbf{u}\in\mathbb{R}^d}\int_{\mathbf{y}\in\mathbb{R}^d} \mathrm{e}^{-\frac{\mathbf{u}^*(A_n)^2\mathbf{u}}{2}}\,\mathbb{1}_{((\Sigma_n)^2\mathbf{y}+\mathbf{u}\,\in\,B)}\,\nu(\mathbf{y})\,d\mathbf{u}\,d\mathbf{y}\\
&= \frac{\det A_n}{(2\pi)^{d/2}}\int_{\mathbf{b}\in B} \mathrm{e}^{-\frac{\mathbf{b}^*(A_n)^2\mathbf{b}}{2}}\,d\mathbf{b}\left(\int_{\mathbf{y}\in\mathbb{R}^d}\mathrm{e}^{\langle\mathbf{b}\,|\,\mathbf{y}\rangle}\,\mathrm{e}^{-\frac{\mathbf{y}^*(\Sigma_n)^2\mathbf{y}}{2}}\,\nu(\mathbf{y})\,d\mathbf{y}\right).
\end{aligned}$$

Since Σ_n goes to zero, the term in parentheses goes by Lebesgue dominated convergence theorem to $\int_{\mathbb{R}^d}\mathrm{e}^{\langle\mathbf{b}\,|\,\mathbf{y}\rangle}\,\nu(\mathbf{y})\,d\mathbf{y} = \psi(\mathbf{b})$. To simplify a bit the discussion, we suppose:

(e) For some fixed symmetric positive definite matrix A, $A_n = \sqrt{t_n}\,A$ with t_n increasing to infinity.

Then, in particular, the term in parentheses goes to $\psi(\mathbf{b})$ in an increasing way, and the convergence is locally uniform in $\mathbf{b}$. One can even suppose $A = I_d$, the general situation following by obvious changes of coordinates. We can then perform a Laplace method on the previous expression, but in a quite unusual multi-dimensional setting. Set

$$b = \inf\{\sqrt{\langle\mathbf{b}\,|\,\mathbf{b}\rangle}\ \mid\ \mathbf{b}\in B\},$$

and suppose $b \neq 0$. For any $\varepsilon > 0$, we split B into $B_{<b+\varepsilon}$, the set of points of B that are of norm $\|b\|$ smaller than $b+\varepsilon$, and $B_{\geq b+\varepsilon}$. We also denote $B_{\min} = \{\mathbf{x}\in\overline{B}\ \mid\ \sqrt{\langle\mathbf{x}\,|\,\mathbf{x}\rangle} = b\}$.

- The integral over $B_{\geq b+\varepsilon}$ is smaller than the integral on the whole complementary of the ball of radius $b+\varepsilon$ and center the origin. Recall that the measure of the $(d-1)$-dimensional regular sphere is $\mathrm{vol}(\mathbb{S}^{d-1}) = 2\,\frac{\pi^{d/2}}{\Gamma(d/2)}$. So, if $I_{\geq b+\varepsilon}$ denotes the part of $\mathbb{P}[\mathbf{X}_n \in t_nB]$ corresponding to the integral over $B_{\geq b+\varepsilon}$, then

$$\begin{aligned}
I_{\geq b+\varepsilon} &\leq \left(\frac{t_n}{2\pi}\right)^{d/2}\int_{\|\mathbf{x}\|\geq b+\varepsilon}\mathrm{e}^{-\frac{t_n\|\mathbf{x}\|^2}{2}}\,\mathrm{e}^{C\|\mathbf{x}\|}\,d\mathbf{x} \leq \frac{2}{\Gamma(d/2)}\left(\frac{t_n}{2}\right)^{d/2}\int_{b+\varepsilon}^{\infty} w^{d-1}\,\mathrm{e}^{-\frac{t_nw^2}{2}}\,\mathrm{e}^{Cw}\,dw\\
&\leq \frac{2\,\mathrm{e}^{\frac{C^2}{2\,t_n}}}{\Gamma(d/2)}\left(\frac{t_n}{2}\right)^{d/2}\int_{b+\varepsilon-\frac{C}{t_n}}^{\infty}\left(x+\frac{C}{t_n}\right)^{d-1}\mathrm{e}^{-\frac{t_nx^2}{2}}\,dx\\
&\leq \frac{2\,\mathrm{e}^{\frac{C^2}{2\,t_n}}}{\Gamma(d/2)}\left(\frac{t_n}{2}\right)^{d/2}\left(1+\frac{C}{t_nb}\right)^{d-1}\int_{b+\varepsilon-\frac{C}{t_n}}^{\infty}x^{d-1}\,\mathrm{e}^{-\frac{t_nx^2}{2}}\,dx\\
&\leq \frac{1}{\Gamma(d/2)}\left(\frac{t_n}{2}\right)^{d/2-1}\left(1+\frac{C}{t_nb}\right)^{d-1}\left(1-\frac{d-2}{t_nb^2}\right)^{-1}\mathrm{e}^{-\frac{t_n(b+\varepsilon)^2}{2}}\,\mathrm{e}^{C(b+\varepsilon)}\,(b+\varepsilon)^{d-2}
\end{aligned}$$

assuming $\frac{C}{t_n} \leq \varepsilon$, and using integration by parts at the end to estimate the Gaussian integral. To fix the ideas, suppose that $t_n b \geq C$ and $t_n b^2 \geq 2(d-2)$, which is surely the case for n big enough. Then,

$$I_{\geq b+\varepsilon} \leq \frac{2}{\Gamma(d/2)}\,(2\,t_n)^{d/2}\,\mathrm{e}^{-\frac{t_n(b+\varepsilon)^2}{2}}\,\mathrm{e}^{C(b+\varepsilon)}\,(b+\varepsilon)^{d-2}. \tag{10}$$

- As for the integral over $B_{<b+\varepsilon}$, we have to use more precisely the topology of B in the neighborhood of the sphere of radius b. Recall that for any non-negative measurable function over $\mathbb{R}^d$, one can make the polar change of coordinates

$$\int_{\mathbb{R}^d} f(\mathbf{x})\,d\mathbf{x} = \int_{r=0}^{\infty}\left(\int_{\mathbb{S}^{d-1}} f(r\mathbf{x})\;d\mu_{\mathbb{S}^{d-1}}(\mathbf{x})\right) r^{d-1}\,dr,$$

where $\mu_{\mathbb{S}^{d-1}}$ is up to a scalar the unique $\mathrm{SO}(\mathbb{R}^d)$-invariant measure on the sphere — the "spherical" Lebesgue measure. Consequently, we can write

$$\begin{aligned}\big(1+o(1)\big)I_{<b+\varepsilon} &= \left(\frac{t_n}{2\pi}\right)^{d/2}\int_{\|\mathbf{b}\|<b+\varepsilon}\mathrm{e}^{-\frac{t_n\,\|\mathbf{b}\|^2}{2}}\,\mathbb{1}_B(\mathbf{b})\,\psi(\mathbf{b})\,d\mathbf{b}\\ &= \left(\frac{t_n}{2\pi}\right)^{d/2}\mathrm{e}^{-\frac{t_nb^2}{2}}\int_{\|\mathbf{c}\|<b+\varepsilon}\mathrm{e}^{t_n\left(\frac{b^2-\|\mathbf{c}\|^2}{2}\right)}\,\mathbb{1}_B(\mathbf{c})\,\psi(\mathbf{c})\,d\mathbf{c}\\ &= \left(\frac{t_n}{2\pi}\right)^{d/2}\mathrm{e}^{-\frac{t_nb^2}{2}}\int_{r=b}^{b+\varepsilon}\mathrm{e}^{t_n\left(\frac{b^2-r^2}{2}\right)}\,r^{d-1}\left(\int_{\mathbb{S}^{d-1}}\mathbb{1}_B(r\mathbf{c})\,\psi(r\mathbf{c})\,d\mu_{\mathbb{S}^{d-1}}(\mathbf{c})\right)dr,\end{aligned}$$

the multiplicative factor $(1+o(1))$ corresponding to the replacement of the integral $\int_{\mathbf{y}\in\mathbb{R}^d}\mathrm{e}^{\langle\mathbf{b}\,|\,\mathbf{y}\rangle}\,\mathrm{e}^{-\frac{\mathbf{y}^*(\Sigma_n)^2\mathbf{y}}{2}}\,\nu(\mathbf{y})\,d\mathbf{y}$ by $\psi(\mathbf{b})$. Suppose for a moment that the integral in parentheses is continuous with respect to r, or at least continuous at $r=b$. Then, up to a multiplicative factor $(1+o(1))$, we can replace it by the constant

$$\frac{1}{b^{d-1}}\int_{B_{\min}}\psi(\mathbf{b})\,d\mu_{\mathbb{S}^{d-1}(0,b)}(\mathbf{b})$$

where $\mu_{\mathbb{S}^{d-1}(0,b)} = \mu_{\text{surface}}$ is the spherical measure on the sphere of radius b that gives for total weight $b^{d-1}\,\mathrm{vol}(\mathbb{S}^{d-1})$. Hence,

$$I_{<b+\varepsilon} = \left(\frac{t_n}{2\pi}\right)^{d/2}\frac{\mathrm{e}^{-\frac{t_nb^2}{2}}}{b^{d-1}}\left(\int_{B_{\min}}\psi(\mathbf{b})\,d\mu_{\text{surface}}(\mathbf{b})\right)\left(\int_{r=b}^{b+\varepsilon}\mathrm{e}^{t_n\left(\frac{b^2-r^2}{2}\right)}\,r^{d-1}\,dr\right)\,\big(1+o(1)\big). \tag{11}$$

Set $\varepsilon=(t_n)^{-2/3}$. In (10), the expansion of $\mathrm{e}^{-\frac{t_n(b+\varepsilon)^2}{2}}$ gives $\mathrm{e}^{-\frac{t_nb^2}{2}}\,\mathrm{e}^{-b\,(t_n)^{1/3}}\big(1+o(1)\big)$, so

$$I_{\geq b+\varepsilon} \leq \frac{2\,\mathrm{e}^{Cb}\,b^{d-2}}{\Gamma(d/2)}\,(2\,t_n)^{d/2}\,\mathrm{e}^{-\frac{t_nb^2}{2}}\,\mathrm{e}^{-b\,(t_n)^{1/3}}\big(1+o(1)\big).$$

On the other hand, since $t_n\varepsilon=(t_n)^{1/3}$ and $t_n\varepsilon^2=(t_n)^{-1/3}$, in Formula (11), one obtains

$$\begin{aligned}I_{<b+\varepsilon} &= \left(\frac{t_n}{2\pi}\right)^{d/2}\mathrm{e}^{-\frac{t_nb^2}{2}}\left(\int_{B_{\min}}\psi(\mathbf{b})\,d\mu_{\text{surface}}(\mathbf{b})\right)\left(\int_{s=0}^{\varepsilon}\mathrm{e}^{-\frac{t_ns^2}{2}}\,\mathrm{e}^{-t_nbs}\,ds\right)\big(1+o(1)\big)\\ &= \frac{(t_n)^{d/2-1}}{(2\pi)^{d/2}}\mathrm{e}^{-\frac{t_nb^2}{2}}\left(\int_{B_{\min}}\psi(\mathbf{b})\,d\mu_{\text{surface}}(\mathbf{b})\right)\left(\int_{s=0}^{+\infty}\mathrm{e}^{-bt}\,dt\right)\big(1+o(1)\big)\\ &= \left(\frac{t_n}{2\pi}\right)^{d/2}\mathrm{e}^{-\frac{t_nb^2}{2}}\left(\int_{B_{\min}}\frac{\psi(\mathbf{b})}{t_n\|\mathbf{b}\|}\,d\mu_{\text{surface}}(\mathbf{b})\right)\big(1+o(1)\big).\end{aligned}$$

The first term is negligible in comparison to the second term, so, one can reasonably conjecture that

$$\limsup_{n\to\infty}\left(\left(\frac{2\pi}{t_n}\right)^{d/2}(t_n b)\,\exp\left(\frac{t_n b^2}{2}\right)\,\mathbb{P}[\mathbf{X}_n\in t_n B]\right)\leq\int_{B_{\min}}\psi(\mathbf{x})\,\mu_{\text{surface}}(d\mathbf{x}).$$

The reason why one has only a inequality with a lim sup and not a equality with the limit is that the spherical integral with B a closed set (which does not cost anything since $B_{\min}=(\overline{B})_{\min}$) is only upper semi-continuous, see the following Lemma. Therefore,

$$\limsup_{r\to b}\left(\int_{\mathbb{S}^{d-1}}\mathbb{1}_B(r\mathbf{c})\,\psi(r\mathbf{c})\,d\mu_{\mathbb{S}^{d-1}}(\mathbf{c})\right)\leq\int_{\mathbb{S}^{d-1}}\mathbb{1}_B(b\mathbf{c})\,\psi(b\mathbf{c})\,d\mu_{\mathbb{S}^{d-1}}(\mathbf{c}),$$

and then the discussion is the same but with inequalities instead of equalities in the estimates of $I_{<b+\varepsilon}$.

Lemma 4.2. *Let ψ be a non-negative continuous function, and C a closed subset of $\mathbb{R}^d$. The function*

$$F_C : r\mapsto\int_{\mathbb{S}^{d-1}}\psi(r\mathbf{x})\,\mathbb{1}_C(r\mathbf{x})\,d\mu_{\mathbb{S}^{d-1}}(\mathbf{x})$$

is upper semi-continuous on $(0,+\infty)$.

Proof. If instead of $\mathbb{1}_C$ we had a continuous function, then by standard results on integrals of continuous functions in two parameters, the integral F would be continuous in r. However, by using the distance function to C, one can write $\mathbb{1}_C$ as the infimum of continuous functions, and therefore, F_C is itself the infimum of continuous functions, whence upper semi-continuous. ∎

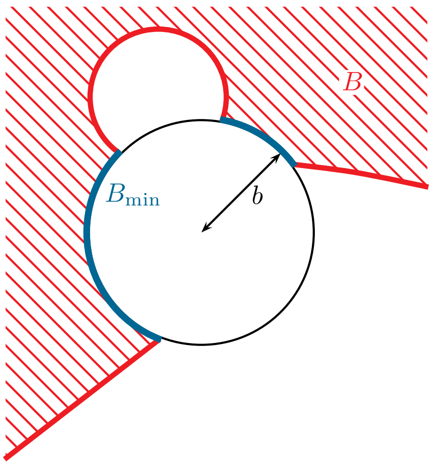


FIGURE 4. In the multi-dimensional mod-Gaussian setting, the precise upper bound is given by the surface integral of ψ on $B_{\min}$, the set of points of B that also lie on the sphere of minimal radius b.

Now working in the more general case of an arbitrary symmetric matrix $A_n=\sqrt{t_n}\,A$, we get the following upper bound:

$$\limsup_{n\to\infty}\left(\left(\frac{2\pi}{t_n}\right)^{d/2}(t_n b)\,\exp\left(\frac{t_n b^2}{2}\right)\,\mathbb{P}[\mathbf{X}_n\in(A_n)^2 B]\right)\leq\int_{B_{\min}}\psi(\mathbf{x})\,\mu_{\text{surface}}(d\mathbf{x})\qquad(12)$$

where $b = \inf\{\sqrt{\langle A\mathbf{b} \mid A\mathbf{b}\rangle} \mid \mathbf{b} \in B\}$; $B_{\min} = \{\mathbf{x} \in \overline{B} \mid \sqrt{\langle A\mathbf{x} \mid A\mathbf{x}\rangle} = b\}$; and μ_{surface} is the surface measure on the A-sphere of radius b obtained by the linear automorphism A from the surface measure on $\mathbb{S}^{d-1}(0,b)$ — in particular it has the same mass, so beware that unless $A = I_d$, this is not the surface measure obtained by restricting the Riemannian structure of the euclidian space $\mathbb{R}^d$ to the A-sphere. The goal of this section is now to prove the estimate (12) assuming only hypotheses (a), (b) and (e); and to give a sufficient condition for the limsup to be a limit.

4.2. **Smoothing techniques and estimates for test functions and domains.** In this section, we assume that 4.1.(a), 4.1.(b) and 4.1.(e) are fulfilled, and we set

$$\mathbf{Y}_n = \frac{\Sigma(\mathbf{X}_n)}{\sqrt{t_n}} = \Sigma_n\, \mathbf{X}_n,$$

where $\Sigma = A^{-1}$. We look for the analogue of Lemma 3.1 in a multi-dimensional setting. Though it might look awkward to work with multi-variables cumulative distribution functions $F(x^{(1)}, \ldots, x^{(d)}) = \mathbb{P}[\mathbf{Y}_n^{(1)} \leq x^{(1)}, \ldots, \mathbf{Y}_n^{(d)} \leq x^{(d)}]$, it seems to be the only way to obtain sharp estimates of the probabilities of hypercubes $\prod_{i=1}^d [a^{(i)}, b^{(i)}]$ under the law μ_n of $\mathbf{Y}_n$. By sharp we mean with a remainder that is a $o(1/\sqrt{t_n})$, and this is needed for our multi-dimensional principle of large dimensions (see §4.3). Nevertheless, one can also give other more general (but less precise) estimates, *e.g.*,

(1) estimates of the difference between $\mathbb{E}[F(\mathbf{Y}_n)]$ and $\mathbb{G}[F] = \int_{\mathbb{R}^d} F(\mathbf{x})\, g(\mathbf{x})\, d\mathbf{x}$ for F in a suitable class of test functions, where $g(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}}\, \mathrm{e}^{-\|\mathbf{x}\|^2/2}$ is the density of a standard d-dimensional Gaussian variable.

(2) and estimates of the difference between $\mu_n(B)$ and $\mathbb{G}(B)$ where B runs over a class of regular domains, for instance, the class of all convex bodies with a sufficiently regular boundary.

Though not needed for our results of large deviations in several dimensions, these new estimates seem of interest in their own right. They should be compared with the approximation results of [BR10], which are also proved in several dimensions, but only in the setting of sums of i.i.d. random variables.

4.2.1. *Estimates for test functions.* A preliminary step consists in smoothing the distribution of the $\mathbf{Y}_n$'s, and to this purpose we shall use standard methods of convolution. Set

$$\Delta_T^{(k)}(x) = \frac{1}{I(k)} \left(\frac{1-\cos Tx}{Tx^2}\right)\left(\frac{\sin Tx}{Tx}\right)^{2k} = \frac{\left(\frac{1-\cos Tx}{Tx^2}\right)\left(\frac{\sin Tx}{Tx}\right)^{2k}}{\int_{-\infty}^{\infty}\left(\frac{1-\cos x}{x^2}\right)\left(\frac{\sin x}{x}\right)^{2k} dx};$$

and in dimension d, $\Delta_{T,\mathbb{R}^d}^{(k)}(\mathbf{x}) = \prod_{i=1}^d \Delta_T^{(k)}(\mathbf{x}^{(i)})$. One defines the convolution of $\Delta_{T,\mathbb{R}^d}^{(k)}$ with a test function F by

$$(\Delta_{T,\mathbb{R}^d}^{(k)} * F)(\mathbf{x}) = \int_{\mathbb{R}^d} F(\mathbf{x}-\mathbf{y})\, \Delta_{T,\mathbb{R}^d}^{(k)}(\mathbf{y})\, d\mathbf{y}.$$

Since $\Delta_{T,\mathbb{R}^d}^{(k)}$ is non-negative and of total mass $\int_{\mathbb{R}^d} \Delta_{T,\mathbb{R}^d}^{(k)}(\mathbf{y})\, d\mathbf{y} = 1$, if μ is a probability measure, then one can define a positive normalized linear form on the space $\mathscr{C}_c(\mathbb{R}^d)$ of continuous compactly supported functions by

$$(\Delta_{T,\mathbb{R}^d}^{(k)} * \mu)(F) = \mu(\Delta_{T,\mathbb{R}^d}^{(k)} * F),$$

and it corresponds by Riesz' representation theorem to a unique probability measure on $\mathbb{R}^d$, the convolution $\mu_T^{(k)} = \Delta^{(k)}_{T,\mathbb{R}^d} * \mu$ of μ with the kernel $\Delta^{(k)}_{T,\mathbb{R}^d}$. This new measure has the following additional properties, which are proved in Appendix 9.1.

Lemma 4.3 (Kernel with compactly supported and smooth Fourier transform)**.**

(1) Suppose F Lipschitz with constant m (w.r.t. the euclidian norm on $\mathbb{R}^d$) and bounded in absolute value by C. Then,

$$|\mu_T^{(k)}(F) - \mu(F)| \leq \frac{27}{2}\, d^{\frac{2k+3}{4k+4}} \left(\frac{C\, m^{2k+1}}{T^{2k+1}}\right)^{\frac{1}{2k+2}}.$$

(2) The Fourier transform of $\Delta^{(k)}_{T,\mathbb{R}^d}$ takes its values in $[0,1]$, is supported by the hypercube $[-(2k+1)T, (2k+1)T]^d$ and is of class $\mathscr{C}^{2k}$ on $\mathbb{R}^d$.

By combining these properties and the hypotheses of mod-convergence, we get:

Lemma 4.4 (Berry-Esseen estimates for test functions)**.** *Let F be a function which is at the same time integrable, bounded and Lipschitz; $C = \|F\|_\infty$ and M a Lipschitz bound for the fluctuations of F. For any $k \geq \lceil \frac{d+1}{2} \rceil$, there exists a constant K_k such that*

$$\begin{aligned}|\mathbb{E}[F(\mathbf{Y}_n)] - \mathbb{G}(F))| &\leq \frac{1}{(t_n)^{1/2}} \left(27(2k+1)\sqrt{d}\,(CM^{2k+1})^{\frac{1}{2k+2}}(dt_n)^{\frac{1}{4k+4}} + CK_k\right)\\ &= O_{M,C,k}\left(\frac{1}{(t_n)^{\frac{1}{2}-\frac{1}{4k+4}}}\right),\end{aligned}$$

and K_k depends only on k, on the speed of convergence of the renormalized characteristic functions on the compact set $\Sigma([-\mathrm{i},\mathrm{i}]^d)$, and on the behavior of ψ on this compact set.

Proof. Fix $k \geq \lceil \frac{d+1}{2} \rceil$. If μ_n is the law of $\mathbf{Y}_n$, then

$$|\mu_n(F) - \mathbb{G}(F)| \leq |\mu_n(F) - \mu^{(k)}_{n,T}(F)| + |\mu_{n,T}(F)^{(k)} - \mathbb{G}^{(k)}_T(F)| + |\mathbb{G}^{(k)}_T(F) - \mathbb{G}(F)| = \alpha + \beta + \gamma.$$

- For α and γ, a correct bound is given by the property (1) of the previous lemma, and we cannot hope for much better since μ_n might be a singular measure (*e.g.* without density). Hence,

$$\alpha + \gamma \leq 27\, d^{\frac{2k+3}{4k+4}} \left(\frac{C\, M^{2k+1}}{T^{2k+1}}\right)^{\frac{1}{2k+2}}.$$

- For β, we use Parseval's theorem. Suppose first that F is in $\mathrm{L}^2(\mathbb{R}^d, d\mathbf{x})$; then,

$$\begin{aligned}|\mu^{(k)}_{n,T}(F) - \mathbb{G}^{(k)}_T(F)| &= \left|\int_{\mathbb{R}^d} F(\mathbf{x}) \left(\frac{d\mu^{(k)}_{n,T}(\mathbf{x})}{d\mathbf{x}} - \frac{d\mathbb{G}^{(k)}_T(\mathbf{x})}{d\mathbf{x}}\right) d\mathbf{x}\right|\\ &= \frac{1}{(2\pi)^d}\left|\int_{\mathbb{R}^d} \widehat{F}(\mathbf{y})\, \widehat{\Delta^{(k)}_{T,\mathbb{R}^d}}(\mathbf{y}) \left(\widehat{\mu_n}(\mathbf{y}) - \widehat{G}(\mathbf{y})\right) d\mathbf{y}\right|.\end{aligned}$$

The identity still holds for $F \in \mathrm{L}^1(\mathbb{R}^d, d\mathbf{x})$, because $\mathrm{L}^1(\mathbb{R}^d, d\mathbf{x}) \cap \mathrm{L}^2(\mathbb{R}^d, d\mathbf{x})$ is dense in $\mathrm{L}^1(\mathbb{R}^d, d\mathbf{x})$. Denote $\psi_n(\mathbf{z})$ the left-hand side of the Equation appearing in

§4.1.(a). With $T = \Theta\sqrt{t_n}$ and $\mathbf{y} \in [-(2k+1)T, (2k+1)T]^d$, one has

$$\widehat{\mu_n}(\mathbf{y}) - \widehat{G}(\mathbf{y}) = (\psi_n(\mathrm{i}\Sigma_n\mathbf{y}) - 1)\,\mathrm{e}^{-\frac{\|\mathbf{y}\|^2}{2}} = \int_0^1 \langle \psi_n'(t\,\mathrm{i}\Sigma_n\mathbf{y}) \,|\, \mathrm{i}\Sigma_n\mathbf{y}\rangle\;\mathrm{e}^{-\frac{\|\mathbf{y}\|^2}{2}}\,dt$$
$$= \frac{1}{\sqrt{t_n}}\int_0^1 \langle \Sigma\psi_n'(t\,\mathrm{i}\Sigma_n\mathbf{y}) \,|\, \mathrm{i}\mathbf{y}\rangle\;\mathrm{e}^{-\frac{\|\mathbf{y}\|^2}{2}}\,dt,$$

and therefore,

$$|\mu_{n,T}^{(k)}(F) - \mathbb{G}_T^{(k)}(F)| \leq \frac{1}{(2\pi)^d\sqrt{t_n}}\int_0^1 \left|\int_{\mathbb{R}^d} \widehat{F}(\mathbf{y})\,\widehat{\Delta_{T,\mathbb{R}^d}^{(k)}}(\mathbf{y})\;\langle \Sigma\psi_n'(t\,\mathrm{i}\Sigma_n\mathbf{y}) \,|\, \mathrm{i}\mathbf{y}\rangle\;\mathrm{e}^{-\frac{\|\mathbf{y}\|^2}{2}}\,d\mathbf{y}\right| dt$$
$$\leq \frac{1}{\sqrt{t_n}}\int_0^1 \left|\int_{\mathbb{R}^d} F(\mathbf{x})\,G_{n,t}(\mathbf{x})\,d\mathbf{x}\right| dt \leq \frac{\|F\|_\infty}{\sqrt{t_n}}\int_0^1\int_{\mathbb{R}^d} |G_{n,t}(\mathbf{x})|\,d\mathbf{x}\,dt$$

where $G_{n,t}$ is the inverse Fourier transform of $\widehat{\Delta_{T,\mathbb{R}^d}^{(k)}}(\mathbf{y})\;\langle \Sigma\psi_n'(t\,\mathrm{i}\Sigma_n\mathbf{y}) \,|\, \mathrm{i}\mathbf{y}\rangle\;\mathrm{e}^{-\frac{\|\mathbf{y}\|^2}{2}}$, that is to say that

$$G_{n,t}(\mathbf{x}) = \int_{\mathbb{R}^d} \widehat{\Delta_{T,\mathbb{R}^d}^{(k)}}(\mathbf{y})\;\langle \Sigma\psi_n'(t\,\mathrm{i}\Sigma_n\mathbf{y}) \,|\, \mathrm{i}\mathbf{y}\rangle\;\mathrm{e}^{-\frac{\|\mathbf{y}\|^2}{2} - \langle \mathrm{i}\mathbf{x}\,|\,\mathbf{y}\rangle}\,d\mathbf{y}.$$

However,

(1) $\widehat{\Delta_{T,\mathbb{R}^d}^{(k)}}$ is of class $\mathscr{C}^{d+1}$ on $\mathbb{R}^d$ and vanishes outside $[-(2k+1)T, (2k+1)T]^d$;
(2) by the hypothesis of mod-Gaussian convergence, $\psi_n'(t\,\mathrm{i}\Sigma_n\mathbf{y})$ and all its derivatives up to order $(d+1)$ are uniformly bounded on $[-(2k+1)T, (2k+1)T]^d$ (actually one does not need here the convergence of these quantities);
(3) and $\mathbf{y}\,\mathrm{e}^{-\frac{\|\mathbf{y}\|^2}{2}}$ is a Schwartz function on $\mathbb{R}^d$.

Therefore,

$$\sup_{t\in[0,1],\ n\in\mathbb{N},\ |\alpha|\leq d+1,\ l\leq d+1} \left\|(1+\|\mathbf{y}\|)^l\,D^\alpha H_{n,t}(\mathbf{y})\right\|_\infty < \infty,$$

so by $(d+1)$ integration by parts, $|G_{n,t}(\mathbf{x})|$ is bounded by $L\,(1+\|\mathbf{x}\|)^{-d-1}$, with L constant that only depends on k and on the behavior of the ψ_n's and their derivatives up to order $(d+2)$ on $\Sigma([-(2k+1)\mathrm{i}\Theta, (2k+1)\mathrm{i}\Theta]^d)$. Since $(1+\|\mathbf{x}\|)^{-d-1}$ is integrable on $\mathbb{R}^d$, there is another finite constant K with the same properties of dependence and such that

$$\sup_{t\in[0,1],\ n\in\mathbb{N}} \left(\int_{\mathbb{R}^d} |G_{n,t}(\mathbf{x})|\,d\mathbf{x}\right),$$

so $\beta \leq \frac{CK}{\sqrt{t_n}}$.

To fix the ideas, set $\Theta = 1/(2k+1)$, and denote K_k the corresponding constant, which only depends on k and on the behavior of ψ and the ψ_n's on $\Sigma([-\mathrm{i},\mathrm{i}]^d)$ (more precisely it suffices to have bound on their derivatives up to order $d+2$). Then,

$$|\mathbb{E}[F(\mathbf{Y}_n)] - \mathbb{G}(F))| \leq 27(2k+1)\,d^{\frac{2k+3}{4k+4}}\,\frac{(C\,M^{2k+1})^{\frac{1}{2k+2}}}{(t_n)^{\frac{2k+1}{4k+4}}} + \frac{C\,K_k}{(t_n)^{1/2}}.$$

In particular, $|\mathbb{E}[F(\mathbf{Y}_n)] - \mathbb{G}(F))| = O\left(\frac{1}{(t_n)^{1/2-\varepsilon}}\right)$ for any $\varepsilon > 0$, with a constant that only depends on ε and M, C. □

4.2.2. *Estimates for Gaussian regular domains.* Starting from the previous Lemma, one can estimate the difference of probabilities $|\mu_n(B) - \mathbb{G}(B)|$ for B domain in $\mathbb{R}^d$, but one needs B to be "sufficiently regular". The right condition to impose is the following. Denote $B^\varepsilon = \{\mathbf{x} \mid d(\mathbf{x}, B) \leq \varepsilon\}$ for $\varepsilon > 0$, and $B^{-\varepsilon} = ((B^c)^\varepsilon)^c$. When B is a convex body, these are the Minkowski sum and the Minkowski difference of B with the ball $B_{(\mathbf{0},\varepsilon)}$.

Definition 4.5. *A domain $B \subset \mathbb{R}^d$ will be called Gaussian regular with constant G if for all $\varepsilon > 0$,*

$$\frac{1}{(2\pi)^{d/2}} \int_{B^\varepsilon \setminus B} \mathrm{e}^{-\frac{\|\mathbf{x}\|^2}{2}}\, d\mathbf{x} \leq G\varepsilon \qquad ; \qquad \frac{1}{(2\pi)^{d/2}} \int_{B \setminus B^{-\varepsilon}} \mathrm{e}^{-\frac{\|\mathbf{x}\|^2}{2}}\, d\mathbf{x} \leq G\varepsilon.$$

Example 4.6. Consider an hypercube, possibly with infinite bounds and possibly rotated by some $u \in \mathrm{SO}(d, \mathbb{R})$:

$$B = u\left(\prod_{i=1}^{d} [a^{(i)}, b^{(i)}]\right), \quad u \in \mathrm{SO}(d, \mathbb{R}), \quad a^{(i)}, b^{(i)} \in \overline{\mathbb{R}}.$$

As shown by Figure 5, the first integral in Definition 4.5 is bounded by

$$\frac{2d}{(2\pi)^{d/2}} \sup_{a \in \mathbb{R}} \left(\int_a^{a+\varepsilon} \mathrm{e}^{-\frac{x^2}{2}}\, dx \int_{\mathbb{R}^{d-1}} \mathrm{e}^{-\frac{(\mathbf{x}^{(2)})^2 + \cdots + (\mathbf{x}^{(d)})^2}{2}}\, d\mathbf{x} \right) \leq \sqrt{\frac{2}{\pi}}\, d\, \varepsilon.$$

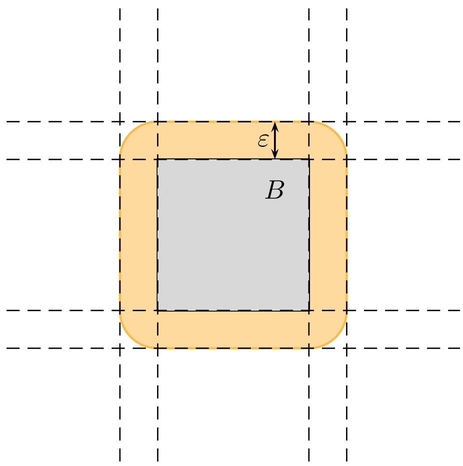


FIGURE 5. For hypercubes, the Gaussian mass of the ε-boundary is bounded by the mass of $2d$ bands $H \times [0, \varepsilon]$ with H hyperplane, and therefore by a constant times ε.

The same bound holds for the second integral, so, hypercubes are Gaussian regular with constant $\sqrt{2/\pi}\, d$. More generally, the same discussion shows that any convex polytope with F faces on its boundary is Gaussian regular with constant $\frac{F}{\sqrt{2\pi}}$.

Example 4.7. A more general set of domains for which our results will apply consists in convex bodies, the asymptotics of Laplace or Fourier transforms of the indicators of convex bodies being a well-known problem (see *e.g.* [Ste93]). The Gaussian regularity of convex sets follows for instance from [BR10, Chapter I, Theorem 3.1]; for completeness we give in Appendix 9.2 a quite shorter proof of this result.

Theorem 4.8 (Berry-Esseen estimates for probabilities of Gaussian regular domains)**.** *Fix $\varepsilon > 0$. For $n \geq n_\varepsilon$ big enough and for all Gaussian regular domain B with constant G,*

$$|\mu_n(B) - \mathbb{G}(B)| \leq \frac{\sqrt{G}}{(t_n)^{\frac{1}{4}-\varepsilon}}.$$

Here n_ε depends on d, ε, Σ and on the behavior of the ψ_n's on $\Sigma([-\mathrm{i},\mathrm{i}]^d)$, but is independent from B and G.

Proof. Fix $k \geq \lceil \frac{d+1}{2} \rceil$. For $M > 0$, set

$$\psi_{B,M}(\cdot) = (1 - M\,d(\cdot, B))_+ \quad ; \quad \phi_{B,M} = 1 - (1 - M\,d(\cdot, B^{\mathrm{c}}))_+.$$

These are Lipschitz functions with constant M and bound $C = 1$. One has

$$\mu_n(B) - \mathbb{G}(B) \leq \mu_n(\psi_{B,M}) - \mathbb{G}(B) \leq |\mu_n(\psi_{B,M}) - \mathbb{G}(\psi_{B,M})| + \int_{B^{1/M}\setminus B} \frac{\mathrm{e}^{-\frac{\|\mathbf{x}\|^2}{2}}}{(2\pi)^{d/2}}\,d\mathbf{x}$$

$$\mu_n(B) - \mathbb{G}(B) \geq \mu_n(\phi_{B,M}) - \mathbb{G}(B) \geq -|\mu_n(\phi_{B,M}) - \mathbb{G}(\phi_{B,M})| - \int_{B\setminus B^{-1/M}} \frac{\mathrm{e}^{-\frac{\|\mathbf{x}\|^2}{2}}}{(2\pi)^{d/2}}\,d\mathbf{x}.$$

By hypothesis, the Gaussian integrals are bounded by $\frac{G}{M}$, and we can use Lemma 4.4 in order to bound the remaining terms:

$$|\mu_n(B) - \mathbb{G}(B)| \leq \frac{27\,(2k+1)\,d^{\frac{2k+3}{4k+4}}\,M^{\frac{2k+1}{2k+2}}}{(t_n)^{\frac{2k+1}{4k+4}}} + \frac{G}{M} + \frac{K_k}{(t_n)^{1/2}},$$

for any $k \geq \lceil \frac{d+1}{2} \rceil$. It suffices now to choose the best M, which is

$$M = \left(\frac{(2k+2)\,G}{27\,(2k+1)^2}\,\frac{(t_n)^{\frac{2k+1}{4k+4}}}{d^{\frac{2k+3}{4k+4}}} \right)^{\frac{2k+2}{4k+3}},$$

and yields the inequality

$$|\mu_n(B) - \mathbb{G}(B)| \leq 14.08\,\sqrt{2k+1}\;\frac{G^{\frac{2k+1}{4k+3}}\,d^{\frac{(2k+2)(2k+3)}{(4k+3)(4k+4)}}}{(t_n)^{\frac{(2k+1)(2k+2)}{(4k+3)(4k+4)}}} + \frac{K_k}{(t_n)^{1/2}},$$

the constant 14.08 corresponding to the largest value of $\frac{1}{\sqrt{2k+1}}\,\frac{4k+3}{2k+1}\left(\frac{27(2k+1)^2}{2k+2}\right)^{\frac{2k+2}{4k+3}}$ (when $k = 1$). Therefore, for $n \geq n_k$ with n_k depending only on k and on the behavior of the ψ_n's on $\Sigma([-\mathrm{i},\mathrm{i}]^d)$,

$$|\mu_n(B) - \mathbb{G}(B)| \leq \frac{29}{2}\,\frac{\sqrt{G\,(2k+1)\,d}}{(t_n)^{\frac{(2k+1)(2k+2)}{(4k+3)(4k+4)}}},$$

making again some simplifications on the constants involved in this upper bound. Since $d \leq 2k-1$, setting $\varepsilon = \frac{1}{16k}$, this gives

$$|\mu_n(B) - \mathbb{G}(B)| \leq \frac{29}{16\varepsilon}\,\frac{\sqrt{G}}{(t_n)^{\frac{1}{4}-\varepsilon}\,(t_n)^{\frac{1}{16k}-\frac{1}{4(4k+3)}}}.$$

For n big enough, $\frac{16\varepsilon}{29}\,(t_n)^{\frac{1}{16k}-\frac{1}{4(4k+3)}}$ is bigger than 1, so, possibly raising the value of $n_k = n_\varepsilon$, we end up with

$$|\mu_n(B) - \mathbb{G}(B)| \leq \frac{\sqrt{G}}{(t_n)^{\frac{1}{4}-\varepsilon}},$$

and this holds for any Gaussian regular domain. □

4.2.3. *Estimates for hypercubes and polytopes.* The previous result shows that for any $\varepsilon > 0$, there exists an integer n_ε such that

$$|\mu_n(C) - \mathbb{G}(C)| \leq \frac{1}{(t_n)^{\frac{1}{4}-\varepsilon}}$$

uniformly over hypercubes C in $\mathbb{R}^d$. However, in dimension $d = 1$, we already know that the difference is in fact a uniform $O((t_n)^{-1/2})$, so we might reasonably conjecture that the previous estimates are not optimal in this case. Indeed, the following is true:

Theorem 4.9 (Berry-Esseen estimates for probabilities of polytopes)**.** *For every d-dimensional hypercube $C = \{\sum_{i=1}^d t^{(i)} v_i,\ \ t^{(i)} \in [a^{(i)}, b^{(i)}]\}$, with the v_i's not necessarily orthogonal, and possibly with infinite bounds,*

$$\mu_n(C) = \frac{1}{(2\pi)^{d/2}} \left(\int_C \mathrm{e}^{-\frac{\|\mathbf{x}\|^2}{2}} \left(1 + \langle \Sigma_n \psi'(\mathbf{0}) \mid \mathbf{x} \rangle\right) d\mathbf{x} \right) + o\left(\frac{1}{\sqrt{t_n}} \right),$$

with a remainder uniform on hypercubes. The same uniform estimate holds over convex polytopes with a bounded number of faces F.

Remark 4.10. The estimate on multi-variables cumulative distribution functions can in turn be used to get better estimates on test functions, but with much stronger hypotheses than in §4.2.1. Hence, if F is a Schwartz function (or at least Schwartz "up to order d" in the same sense as in the proof of Lemma 4.4), then by integration by parts one can:

- write $\mu_n(F)$ as the integral of the d-th derivative of F against the cumulative distribution function of $\mathbf{Y}_n$;
- replace this cumulative distribution function by its Gaussian approximation;
- and undo the d integrations by parts to obtain $\mathbb{G}(f)$ up to a $O((t_n)^{-1/2})$, instead of a $O((t_n)^{-1/2+\varepsilon})$.

However we need F to be much smoother than before, and this only removes the $-\varepsilon$ in the exponent of the estimates.

Proof. The main difference between the polytope case and the previous computations is the possibility to use the *monotony* of the cumulative distribution function

$$F(x^{(1)}, \ldots, x^{(d)}) = \mathbb{P}\left[\mathbf{Y}_n = \sum_{i=1}^d y^{(i)} v_i,\ \ y^{(i)} \leq x^{(i)} \right] = \mu_n(C(x^{(1)}, \ldots, x^{(d)}))$$

with respect to all of its parameters. Denote

$$G(x^{(1)}, \ldots, x^{(d)}) = \frac{1}{(2\pi)^{d/2}} \int_{C(x^{(1)},\ldots,x^{(d)})} \mathrm{e}^{-\frac{\|\mathbf{x}\|^2}{2}} \left(1 + \langle \mathbf{v} \mid \mathbf{x} \rangle\right) d\mathbf{x};$$

$\delta(x^{(1)}, \ldots, x^{(d)}) = (F - G)(x^{(1)}, \ldots, x^{(d)})$ and $\eta = \sup_{x^{(1)},\ldots,x^{(d)}} |\delta(x^{(1)}, \ldots, x^{(d)})|$. The d-dimensional analogue of [Fel71, Lemma XVI.3.1] yields

$$\eta \leq 2\eta_T + \frac{M}{T},$$

with

$$\begin{aligned}\eta_T &= \sup_{x^{(1)},\dots,x^{(d)}} |\delta_T(x^{(1)},\dots,x^{(d)})| \\ &= \sup_{x^{(1)},\dots,x^{(d)}} |(\Delta^{(0)}_{T,\mathbb{R}^d} * F)(x^{(1)},\dots,x^{(d)}) - (\Delta^{(0)}_{T,\mathbb{R}^d} * G)(x^{(1)},\dots,x^{(d)})|\end{aligned}$$

where M is a multiple of a Lipschitz bound for G, that is to say a constant depending only on $\mathbf{v}$. The monotony of F plays here a crucial role: indeed, Feller's proof uses the fact that if $(x^{(1)},\dots,x^{(d)})$ are parameters such that $\eta = \delta(x^{(1)},\dots,x^{(d)})$, then for a certain vector $\mathbf{m} \in \mathbb{R}^d$,

$$\delta(\widetilde{x}^{(1)}+e^{(1)},\dots,\widetilde{x}^{(d)}+e^{(d)}) \geq F(x^{(1)},\dots,x^{(d)}) - G(\widetilde{x}^{(1)}+e^{(1)},\dots,\widetilde{x}^{(d)}+e^{(d)}) \geq \frac{\eta}{2} + \langle \mathbf{m} \,|\, \mathbf{e}\rangle$$

in a neighborhood $\prod_{i=1}^d [\widetilde{x}^{(i)}-\varepsilon^{(i)}, \widetilde{x}^{(i)}+\varepsilon^{(i)}]$ of $(\widetilde{x}^{(1)},\dots,\widetilde{x}^{(d)}) = (x^{(1)}+\varepsilon^{(1)},\dots,x^{(d)}+\varepsilon^{(d)})$. Now, the same arguments as in Lemma 3.1 yield the uniform bound of the theorem for infinite hypercubes $C(x^{(1)},\dots,x^{(d)})$ — notice that the term $\eta'''(\mathbf{0})$ disappears here since $\eta''' = 0$. Then, for any convex polytope C bounded by F faces (including bounded hypercubes with $F = 2d$), one can use inclusion-exclusion to write $\mathbb{1}_C$ as a linear combination of 2^F cumulative distribution functions:

$$\mathbb{1}_C = \prod_{f \in \mathrm{faces}(C)} \mathbb{1}_{f^+} = \prod_{f \in F(C)} (1 - \mathbb{1}_{f^-}) = \sum_{A \subset F(C)} \pm \mathbb{1}_{C(A)}, \tag{13}$$

where $\mathbb{1}_{f^+}$ and $\mathbb{1}_{f^-}$ denote the indicator functions of the two half-spaces determined by the affine hyperplane f, with the $+$ side containing C; and for A subset of the set of faces $F(C)$, $C(A)$ is the corresponding infinite hypercube $\bigcap_{f\in A} f^-$. Since the bound $o((t_n)^{-1/2})$ for $|\mu_n(C) - G(C)|$ is uniform over infinite hypercubes, we get from (13) the same uniform bound for convex polytopes with a bounded number of faces. ☐

4.3. **Estimates of probabilities of Borel sets in a multi-dimensional setting.** We are now ready to prove the large deviations principle (12); again, we suppose $A = \Sigma = I_d$, as this only amounts to easy changes of variables. Fix $\mathbf{x} \neq \mathbf{0}$, and as before, introduce a random variable $\widetilde{\mathbf{X}}_n$ with probability

$$\mathbb{Q}_n[d\mathbf{y}] = \frac{\exp(\langle \mathbf{y} \,|\, \mathbf{x}\rangle)}{\phi_n(\mathbf{x})}\, \mathbb{P}_n[d\mathbf{y}],$$

where $\mathbb{P}_n$ is the law of $\mathbf{X}_n$. The characteristic function of $\widetilde{\mathbf{X}}_n$ is $\frac{\phi_n(\mathbf{z}+\mathbf{x})}{\phi_n(\mathbf{x})}$, so modulo

$$\exp\left(\frac{t_n \|\mathbf{z}\|^2}{2} + \langle t_n \mathbf{x} \,|\, \mathbf{z}\rangle\right),$$

which is the Laplace transform of a Gaussian random variable with mean $t_n\mathbf{x}$ and covariance matrix $t_n\, I_d$, it converges locally uniformly to $\frac{\psi(\mathbf{z}+\mathbf{x})}{\psi(\mathbf{x})}$. Notice that if $\mathbf{x}$ stays in a compact subset of $\mathbb{R}^d$, then the bounds measuring this convergence can be made independent of $\mathbf{x}$, so the previous theory applies locally uniformly with respect to $\mathbf{x}$. In particular, for any convex polytope C,

$$\left|\mathbb{Q}\left[\frac{\widetilde{\mathbf{X}}_n - t_n\,\mathbf{x}}{\sqrt{t_n}} \in C\right] - \mathbb{G}_{\psi'(\mathbf{x})}(C)\right| = o\left(\frac{1}{(t_n)^{1/2}}\right), \tag{14}$$

where $\mathbb{G}_{\psi'(\mathbf{x})}$ is the signed measure with density $\left(1 + \frac{\langle \psi'(\mathbf{x}) \,|\, \mathbf{y}\rangle}{\sqrt{t_n}\,\psi(\mathbf{x})}\right) \mathrm{e}^{-\frac{\|\mathbf{y}\|^2}{2}}\, d\mathbf{y}$.

We fix a polytope neighborhood B of $\mathbf{0}$ in the $(d-1)$ vector space orthogonal to $\mathbf{x}$, and we consider the convex polytopes

$$C(\mathbf{0},B) = \{(\lambda-1)\mathbf{x}+\lambda\mathbf{b},\ \mathbf{b}\in B,\ \lambda\geq 1\};$$
$$C(\mathbf{x},B) = \mathbf{x}+C(\mathbf{0},B) = \{\lambda(\mathbf{x}+\mathbf{b}),\ \mathbf{b}\in B,\ \lambda\geq 1\},$$

see the following Figure.

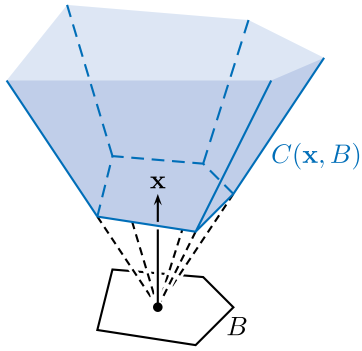


FIGURE 6. For cones $C(\mathbf{x},B)$ with polytope basis B, one can approximate the probability $\mathbb{P}[\mathbf{X}_n \in t_n\, C]$ by the conjectured upper bound, and this locally uniformly in $\mathbf{x}$.

Denote μ_n the law of $\frac{\widetilde{\mathbf{X}}_n - t_n\mathbf{x}}{\sqrt{t_n}}$, and suppose $\|\mathbf{x}\|=x$. One has

$$\begin{aligned}\mathbb{P}[\mathbf{X}_n \in t_n\, C(\mathbf{x},B)] &= \phi_n(\mathbf{x}) \int_{\mathbb{R}^d} 1_{\mathbf{y}\in t_n\, C(\mathbf{x},B)}\, \exp(-\langle \mathbf{y}\,|\,\mathbf{x}\rangle)\, \mathbb{Q}_n(d\mathbf{y})\\ &= \psi_n(\mathbf{x})\, \exp\left(-\frac{t_n x^2}{2}\right) \int_{\mathbb{R}^d} 1_{\mathbf{u}\in\sqrt{t_n}\, C(\mathbf{0},B)}\, \exp(-\sqrt{t_n}\,\langle \mathbf{u}\,|\,\mathbf{x}\rangle)\, \mu_n(d\mathbf{u}).\end{aligned}$$

By slicing the cone $C(\mathbf{0},B)$ orthogonally to the direction $\mathbf{x}$, we see that the integral is also equal to

$$I = \int_0^\infty t_n x^2\, \exp(-\lambda\, t_n x^2)\, \mu_n(C_{n,\lambda})\, d\lambda,$$

where $C_{n,\lambda}$ is the part of $\sqrt{t_n}\, C(\mathbf{0},B)$ consisting of elements $(\mu\mathbf{x}+(\mu+\sqrt{t_n})\mathbf{b})$ with $0\leq\mu\leq\lambda$. We can now replace $\mu_n(C_{n,\lambda})$ by its approximation (14), and as in dimension 1, it suffices to take the Gaussian mass, as this is the only part that will give something bigger than $o(1/\sqrt{t_n})$. So,

$$\begin{aligned}I &= \int_0^\infty t_n x^2\, \exp(-\lambda\, t_n x^2)\left(\mathbb{G}(C_{n,\lambda}) + o\left(\frac{1}{\sqrt{t_n}}\right)\right) d\lambda\\ &\simeq \frac{x}{(2\pi)^{d/2}} \int_0^\infty (\lambda+\sqrt{t_n})^{d-1} \left(\int_B \exp\left(-\lambda\, t_n x^2 - \frac{\lambda^2 x^2 + (\lambda+\sqrt{t_n})^2\|\mathbf{b}\|^2}{2}\right) d\mathbf{b}\right) d\lambda\end{aligned}$$

up to $o((t_n)^{-1/2})$ on the last line.

Let ε be function going to 0 as its argument goes to zero, and such that the uniform remainder in the previous formula always satisfies

$$\left|o\left(\frac{1}{\sqrt{t_n}}\right)\right| \leq \frac{1}{\sqrt{t_n}}\,\varepsilon^3\left(\frac{1}{\sqrt{t_n}}\right).$$

We suppose that $B = \frac{S}{\sqrt{t_n}}$ with S polytope of surface measure of order $\frac{1}{\psi_n(\mathbf{x})}\,\varepsilon\Big(\frac{1}{\sqrt{t_n}}\Big)$; here by of order we mean equal up to a positive multiplicative constant bounded from below and from above. Under this assumption, the asymptotics of the integral are

$$\left|I - \frac{(t_n)^{d/2-1}}{(2\pi)^{d/2}\,x}\,\mu_{\mathrm{surface}}(B)\right| \leq \frac{1}{\sqrt{t_n}}\,\varepsilon^2\left(\frac{1}{\sqrt{t_n}}\right),$$

and the estimate is of order $(t_n)^{d/2-1-(d-1)/2}\,\mu(S)$, that is to say of order $\frac{1}{\sqrt{t_n}}\,\varepsilon\Big(\frac{1}{\sqrt{t_n}}\Big)$, so larger than the remainder. Indeed, in the integral over B, $\|\mathbf{b}\|$ is always a $o(1)$ by assumption on the size of the surface, and then one only has to compute a one dimensional Gaussian integral, which is done as in the proof of Theorem 3.2. So:

Lemma 4.11. *Fix $\mathbf{x} \neq \mathbf{0}$ and a $(d-1)$-dimensional polytope S orthogonal to $\mathbf{x}$ and of surface measure of order $\frac{1}{\psi_n(\mathbf{x})}\,\varepsilon\Big(\frac{1}{\sqrt{t_n}}\Big)$, with ε chosen as before. One has*

$$\exp\left(\frac{t_n\,x^2}{2}\right)\,\mathbb{P}\left[\mathbf{X}_n \in t_n\,C\left(\mathbf{x}, \frac{S}{\sqrt{t_n}}\right)\right] = \frac{(t_n)^{d/2-1}\,\psi_n(\mathbf{x})}{(2\pi)^{d/2}\,x}\;\mu_{\mathrm{surface}}\left(\frac{S}{\sqrt{t_n}}\right)$$

up to a remainder smaller than $\frac{1}{\sqrt{t_n}}\,\varepsilon^2\Big(\frac{1}{\sqrt{t_n}}\Big)$.

This leads finally to the analogue of Theorems 3.2 and 3.9 in several dimensions:

Theorem 4.12. *Fix a measurable part $\mathcal{S}$ of the A-sphere $\{\mathbf{x},\ \|A\mathbf{x}\| = b\}$, $b > 0$; and denote $\mathcal{S}^+ = [1,+\infty)\,\mathcal{S}$. Locally uniformly in b,*

$$\mathbb{P}[\mathbf{X}_n \in (A_n)^2\,\mathcal{S}^+] = \left(\frac{t_n}{2\pi}\right)^{\frac{d}{2}} \exp\left(-\frac{t_n b^2}{2}\right)\left(\int_{\mathcal{S}} \frac{\psi(\mathbf{x})}{t_n b}\,d\mu_{\mathrm{surface}}(d\mathbf{x})\right)\big(1+o(1)\big).$$

Proof. Supposing again $A = \Sigma = I_d$, it suffices to approximate the surface $\mathcal{S}$ by polytopes $\frac{S}{\sqrt{t_n}}$ with the hypotheses of the previous Lemma (see the next Figure). More precisely, we shall take two kinds of approximations:

- outer approximations $\mathcal{S}_{\mathrm{ext}}$, such that the projection of $\mathcal{S}_{\mathrm{ext}}$ onto the sphere is included into $\mathcal{S}$ and converges in measure to it; the disjoint union of polytopes $\mathcal{S}^+_{\mathrm{ext}}$ is always included into $\mathcal{S}^+$ and $\lim \int_{\mathcal{S}_{\mathrm{ext}}} \psi(\mathbf{b})\,d\mu_{\mathrm{surface}}(\mathbf{b}) = \int_{\mathcal{S}} \psi(\mathbf{b})\,d\mu_{\mathrm{surface}}(\mathbf{b})$.

- inner approximations $\mathcal{S}_{\mathrm{int}}$, such that the projection of $\mathcal{S}_{\mathrm{int}}$ onto the sphere contains $\mathcal{S}$ and converges in measure to it; the disjoint union of polytopes $\mathcal{S}^+_{\mathrm{int}}$ always contains $\mathcal{S}^+$ and $\lim \int_{\mathcal{S}_{\mathrm{int}}} \psi(\mathbf{b})\,d\mu_{\mathrm{surface}}(\mathbf{b}) = \int_{\mathcal{S}} \psi(\mathbf{b})\,d\mu_{\mathrm{surface}}(\mathbf{b})$.

In each case, one needs a number of polytopes of order larger than $\varepsilon^{-1}\,(t_n)^{(d-1)/2}$, so let us take an approximation with $\varepsilon^{-2}\,(t_n)^{(d-1)/2}$ such polytopes. The sum of the remainders will be smaller in order than

$$\frac{\varepsilon^3}{\sqrt{t_n}} \times \frac{1}{\varepsilon^2}\,(t_n)^{\frac{d-1}{2}} = \varepsilon\,(t_n)^{d/2-1} = o((t_n)^{d/2-1}).$$

Then, for the outer approximations,

$$\begin{aligned}\mathbb{P}[\mathbf{X}_n \in (A_n)^2\,\mathcal{S}^+] &\geq \mathbb{P}[\mathbf{X}_n \in (A_n)^2\,\mathcal{S}^+_{\mathrm{ext}}]\\ &\geq \left(\frac{t_n}{2\pi}\right)^{\frac{d}{2}} \exp\left(-\frac{t_n b_+^2}{2}\right)\left(\int_{\mathcal{S}_{\mathrm{ext}}} \frac{\psi(\mathbf{x})}{t_n b_+}\,\mu_{\mathrm{surface}}(d\mathbf{x})\right)\big(1+o(1)\big)\end{aligned}$$

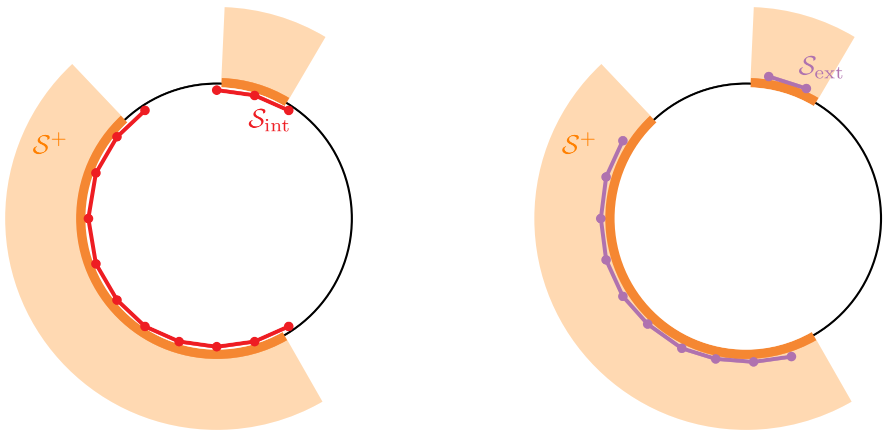


FIGURE 7. The two approximations of a part of the sphere by polytopes.

where b_+ is the norm of the points that are the centers of the polytope faces of the approximation. Since the polytopes are of width $o(1/\sqrt{t_n})$, one can replace b_+ by b without changing the previous formula, which proves the lower bound

$$\liminf_{n\to\infty}\left(\left(\frac{2\pi}{t_n}\right)^{\frac{d}{2}}(t_n b)\,\exp\left(\frac{t_n b^2}{2}\right)\,\mathbb{P}[\mathbf{X}_n\in(A_n)^2B]\right)\geq\int_{\mathcal{S}}\psi(\mathbf{x})\,\mu_{\mathrm{surface}}(d\mathbf{x}).$$

Taking inner approximations, one obtains the same upper bound for the limsup. ☐

Theorem 4.13. *Let B be a Borelian subset of $\mathbb{R}^d$. With the notations of Formula (12),*

$$\limsup_{n\to\infty}\left(\left(\frac{2\pi}{t_n}\right)^{d/2}(t_n b)\,\exp\left(\frac{t_n b^2}{2}\right)\,\mathbb{P}[\mathbf{X}_n\in(A_n)^2B]\right)\leq\int_{B_{\min}}\psi(\mathbf{x})\,\mu_{\mathrm{surface}}(d\mathbf{x})$$

with equality if for instance $B=C$ is closed and the function F_C of Lemma 4.2 is continuous at $r=b$.

Proof. These are now the same arguments as for the toy-model, *cf.* §4.1. ☐

5. First examples

The general results of Sections 2-4 can be applied in many contexts, and the main difficulty is then to prove for each case that one has indeed the estimate on the Laplace transform given by Definition 1.1. An explicit formula for this characteristic function, which is mainly accessible for sums of i.i.d. variables (§5.1), is fortunately not required. Therefore, the development of techniques to obtain mod-ϕ estimates becomes an interesting part of the work. In probabilistic number theory, this will usually be related to the Selberg-Delange method (§5.2), whereas for random combinatorial objets, we will have to combine the methods of §3.2 with some new tools (Sections 6-8). In this section, we detail examples for which the mod-ϕ convergence has already been proved before (*cf.* [JKN11, DKN11]).

5.1. **Sums of independent random variables.** Though the theory of mod-ϕ convergence is meant to be used with complex random variables (statistics of random combinatorial objects, arithmetic properties of random integers, sums of dependent random variables, *etc.*), it already gives interesting results for sums of independent and identically distributed random variables. Let $\mathbf{X}$ be a random variable in $\mathbb{R}^d$ with entire characteristic function, and $(\mathbf{X}_n)_{n\in\mathbb{N}}$ be independent copies of it. The sum $\mathbf{S}_n = \sum_{i=1}^n \mathbf{X}_i$ has characteristic function $\mathrm{e}^{n\langle \mathbf{X}\,|\,\mathbf{z}\rangle}$, so $\mathbf{Z}_n = n^{-1/3}\,\mathbf{S}_n$ converges modulo a Gaussian of mean $n^{2/3}\,\mathbb{E}[\mathbf{X}]$ and variance matrix $n^{1/3}\,\mathrm{Var}[\mathbf{X}]$, with limiting function

$$\psi(\mathbf{z}) = \exp\left(\frac{\kappa^{(3)}(\mathbf{X})(\mathbf{z}^{\otimes 3})}{6}\right).$$

In particular, the following moderate deviation principle holds: assuming $\mathbb{E}[\mathbf{X}] = 0$ and $\mathbf{X}$ truly d-dimensional, *i.e.*, $A = \mathrm{var}[\mathbf{X}] = \mathrm{E}[\mathbf{X}^{\otimes 2}]$ is symmetric positive definite, one has

$$\mathbb{P}\big[\|\varSigma\,\mathbf{S}_n\| \geq n^{2/3} b\big] \simeq \frac{(n^{1/6}\,b)^{d-2}}{(2\pi)^{\frac{d}{2}}}\,\mathrm{e}^{-\frac{(n^{1/6}\,b)^2}{2}}\,\left(\int_{\mathbb{S}^{d-1}(0,1)} \exp\left(\frac{\kappa^{(3)}(\mathbf{X})(\mathbf{z}^{\otimes 3})\,b^3}{6}\right)\mu_{\text{surface}}(d\mathbf{z})\right)$$

where $\varSigma = A^{-1}$. This estimates holds for b of order bigger than $n^{-1/6}$, and up to order $o(n^{1/12})$ according to the discussion of §3.2. If $\mathbf{X}$ is symmetric in law ($\mathbf{X}$ and $-\mathbf{X}$ have same law), then the third cumulant vanishes and one has to look for the fourth cumulant: hence, the random walk $\mathbf{S}_n$ satisfies in this case the moderate deviation principle

$$\mathbb{P}\big[\|\varSigma\,\mathbf{S}_n\| \geq n^{3/4} b\big] \simeq \frac{(n^{1/4}\,b)^{d-2}}{(2\pi)^{\frac{d}{2}}}\,\mathrm{e}^{-\frac{(n^{1/4}\,b)^2}{2}}\,\left(\int_{\mathbb{S}^{d-1}(0,1)} \exp\left(\frac{\kappa^{(4)}(\mathbf{X})(\mathbf{z}^{\otimes 4})\,b^4}{24}\right)\mu_{\text{surface}}(d\mathbf{z})\right)$$

for b of order bigger than $n^{-1/4}$, and up to order $o(n^{1/12})$.

A simple consequence of these multi-dimensional results is the loss of symmetry of the random walks on $\mathbb{Z}^d$ conditioned to be far away from the origin; this loss of symmetry has also been brought out in dimension 2 in the recent paper [Ben13]. Thus, consider the simple 2-dimensional random walk $\mathbf{S}_n = \sum_{i=1}^n \mathbf{X}_i$, where $\mathbf{X}_i = (\pm 1, 0)$ or $(0, \pm 1)$ with probability $1/4$ for each direction. The previous discussion and the values of cumulants

$$\kappa^{(4)}((\Re\mathbf{X})^{\otimes 4}) = \kappa^{(4)}((\Im\mathbf{X})^{\otimes 4}) = \kappa^{(4)}((\Re\mathbf{X})^{\otimes 2}, (\Im\mathbf{X})^{\otimes 2}) = -\frac{1}{4}$$

leads to the following limiting result: if $\mathbf{X}_n = R_n\,\mathrm{e}^{\mathrm{i}\theta_n}$ with $\theta_n \in (0, 2\pi)$, then

$$\lim_{n\to\infty} \mathbb{P}\big[\theta_n \in (\theta_1, \theta_2)\,\big|\,R_n \geq r n^{3/4}\big] = \int_{\theta_1}^{\theta_2} F(r,\theta)\,d\theta$$

with $F(r,\theta) = \dfrac{\exp\left(-\frac{r^4\,(\sin 2\theta)^2}{96}\right)}{\int_0^{2\pi}\exp\left(-\frac{r^4\,(\sin 2\theta)^2}{96}\right)d\theta}$ drawn hereafter.

This function gets concentrated around the two axes of $\mathbb{R}^2$ when $r \to \infty$, whence a loss of symmetry in comparison to the behavior of the 2-dimensional Brownian motion (the scaling limit of the random walk). In dimension $d \geq 3$, one obtains the similar result

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{\mathbf{S}_n}{\|\mathbf{S}_n\|} \in A\,\middle|\,\|\mathbf{S}_n\| \geq r n^{3/4}\right] = K(r)\int_A \exp\left(-\frac{r^4}{12\,d}\sum_{1\leq i<j\leq d} (\mathbf{x}^{(i)}\,\mathbf{x}^{(j)})^2\right)\mu_{\text{surface}}(d\mathbf{x})$$

for any measurable set $A \subset \mathbb{S}^{d-1}(0,1)$. The conditional probability is therefore concentrated around the axes of $\mathbb{R}^d$. All these estimates hold up to $r = o(n^{1/12})$.

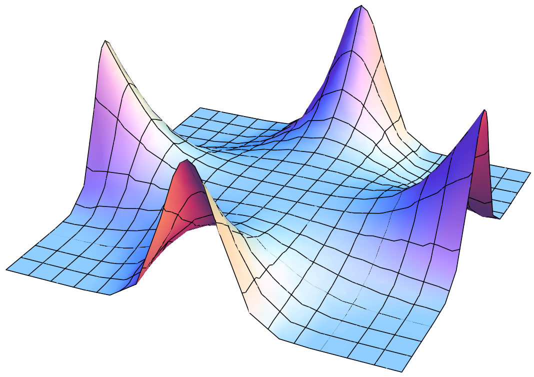


FIGURE 8. The function $F(r,\theta)$ measuring the loss of symmetry of $\mathbf{S}_n$ when conditioned to $\|\mathbf{S}_n\| = O(n^{3/4})$ (using `Mathematica`®).

Another similar setting in which our Theorems apply readily is the Poisson approximation of a sum of Bernoulli variables with "small" parameters p_k. Set $X_n = \sum_{k=1}^{n} \mathcal{B}(p_k)$, where $\mathcal{B}(p)$ is the random variable equal to 1 with probability p and to 0 with probability $1-p$; and where these Bernoulli variables are supposed independent. Under the assumptions $\sum_{k=1}^{\infty} p_k = +\infty$ and $\sum_{k=1}^{\infty}(p_k)^2 < +\infty$, the Laplace transform of X_n writes as

$$\mathbb{E}[\mathrm{e}^{zX_n}] = \prod_{k=1}^{n}(1+p_k(\mathrm{e}^z-1)) = \mathrm{e}^{t_n(\mathrm{e}^z-1)}\left(\prod_{k=1}^{\infty}(1+p_k(\mathrm{e}^z-1))\,\mathrm{e}^{-p_k(\mathrm{e}^z-1)}\right)\big(1+o(1)\big),$$

where $t_n = \sum_{k=1}^{n} p_k \to \infty$. The infinite product $\psi(z)$ is indeed convergent, because

$$(1+p_k(\mathrm{e}^z-1))\,\mathrm{e}^{-p_k(\mathrm{e}^z-1)} = 1 - \frac{(p_k)^2}{2}(\mathrm{e}^z-1) + o((p_k)^2)$$

and $\sum_{k=1}^{\infty}(p_k)^2$ is finite. So, one has mod-Poisson convergence, and by Theorem 2.4 the Poisson approximation holds with precise large deviations

$$\mathbb{P}[X_n \geq (1+\varepsilon)\,t_n] \simeq \frac{\mathrm{e}^{-t_n((1+\varepsilon)\log(1+\varepsilon)-\varepsilon)}}{\sqrt{2\pi t_n}}\,\frac{\sqrt{1+\varepsilon}}{\varepsilon}\,\psi(\varepsilon).$$

5.2. **Logarithmic combinatorial structures.** The previous example is a toy-model for the so-called logarithmic combinatorial structures, see [ABT03, FS90]. Non-trivial examples falling in this framework are the number of cycles of a random permutation (Example 2.12), and the number of distinct prime divisors of a random integer or of a random polynomial over a finite field. Denote $\omega(k)$ the number of distinct prime divisors of an integer k, and ω_n the random variable $\omega(k)$ with k random integer uniformly chosen in $[n] = \{1,2,\ldots,n\}$. The random variable ω_n satisfies the Erdős-Kac central limit theorem (*cf.* [EK40]):

$$\frac{\omega_n - \log\log n}{\sqrt{\log\log n}} \to \mathcal{N}(0,1).$$

Indeed, the Selberg-Delange method (see [Ten95, §2.5] and the next paragraph) yields

$$\mathbb{E}[\mathrm{e}^{z\,\omega_n}] = \mathrm{e}^{(\log\log n+\gamma)(\mathrm{e}^z-1)}\,\Pi_{\mathbb{P}}(\mathrm{e}^z-1)\,\Pi_{\mathbb{N}^*}(\mathrm{e}^z-1)\big(1+o(1)\big),$$

where $\gamma = 0.577...$ is the Euler-Mascheroni constant, and $\Pi_{\mathbb{A}}(x) = \prod_{a\in\mathbb{A}}(1+\frac{x}{a})\,\mathrm{e}^{-\frac{x}{a}}$ for $\mathbb{A}$ part of $\mathbb{N}^* = [\![1,+\infty]\!]$. This mod-Poisson convergence leads now to the principle of large deviations:

$$\mathbb{P}[\omega_n \geq (1+\varepsilon)(\log\log n+\gamma)] \simeq \frac{\mathrm{e}^{-(\log\log n+\gamma)((1+\varepsilon)\log(1+\varepsilon)-\varepsilon)}}{\sqrt{2\pi\log\log n}}\,\frac{\sqrt{1+\varepsilon}}{\varepsilon}\,\Pi_{\mathbb{P}}(\varepsilon)\,\Pi_{\mathbb{N}^*}(\varepsilon).$$

Similarly, denote $\omega_{n,q}$ the number of distinct irreducible divisors of a random monic polynomial of degree n over the finite field $\mathbb{F}_q$ with q elements. It is shown in [KN10, Theorem 6.1] that the characteristic function of $\omega_{n,q}$ has for asymptotics:

$$\mathbb{E}[\mathrm{e}^{z\,\omega_{n,q}}] \simeq \mathrm{e}^{(\log n)(\mathrm{e}^z-1)}\,\frac{1}{\Gamma(\mathrm{e}^z)}\prod_{\pi}\left(1-\frac{1}{q^{\deg\pi}}\right)^{\mathrm{e}^z}\left(1+\frac{\mathrm{e}^z}{q^{\deg\pi}-1}\right),$$

where the product runs over monic irreducible polynomials over $\mathbb{F}_q$. Hence, if $I_{q,d} = \frac{1}{d}\sum_{e|d}\mu(e)\,q^{d/e} = \frac{q^d}{d} + O(q^{d/2})$ is the number of irreducible polynomials of degree d, then

$$\mathbb{E}[\mathrm{e}^{z\,\omega_{n,q}}] \simeq \mathrm{e}^{(\log n)(\mathrm{e}^z-1)}\,\frac{1}{\Gamma(\mathrm{e}^z)}\prod_{d=1}^{\infty}\left(1-\frac{1}{q^d}\right)^{I_{q,d}\,\mathrm{e}^z}\left(1+\frac{\mathrm{e}^z}{q^d-1}\right)^{I_{q,d}},$$

from which one deduces that $\mathbb{P}[\omega_{n,q} \geq (1+\varepsilon)\log n]$ is equivalent to

$$\frac{\mathrm{e}^{-\log n((1+\varepsilon)\log(1+\varepsilon)-\varepsilon)}}{\sqrt{2\pi\log n}}\,\frac{\sqrt{1+\varepsilon}}{\varepsilon\,\Gamma(1+\varepsilon)}\prod_{d=1}^{\infty}\left(\left(1+\frac{\varepsilon}{q^d}\right)\left(1-\frac{1}{q^d}\right)^{\varepsilon}\right)^{I_{q,d}}.$$

In particular, the cardinality q only plays a role in the large deviations of $\omega_{n,q}$, but not in the central limit theorem.

5.3. **Additive arithmetic functions of random integers.** The previous example concerning the number of prime divisors of a random integer can be generalized as follows. Let $f : \mathbb{N} \to \mathbb{R}$ be an arithmetic function with the following properties:

(i) f is additive: $f(mn) = f(m) + f(n)$ if $m \wedge n = 1$;

(ii) $f(p^k) = g(p,k)$ is bounded as a function of prime numbers p and integers k (we shall denote C a bound);

(iii) $f(p) = g(p,1) = 1 + O(\frac{1}{p^\alpha})$ for some $\alpha > 0$.

For the last condition, by dividing f by some constant, one can assume more generally f "almost constant on primes". Denote then

$$F(s,w) = \sum_{n=1}^{\infty}\frac{w^{f(n)}}{n^s}$$

the Dirichlet series of $(w^{f(n)})_{n\in\mathbb{N}}$, which is well-defined for $w \in \mathbb{C}^\times$ and $\sigma = \Re(s) > 1$. One forms the ratio

$$G(s,w) = \frac{F(s,w)}{\zeta(s)^w},$$

where $\zeta(s)$ is the Riemann zeta function $\sum_{n=1}^{\infty}\frac{1}{n^s}$. The choice of complex logarithms implied by the writing $\zeta(s)^w$ is the one of [Ten95, II.5.1]. On the other hand, it is well-known that there exists a constant $\frac{1}{2} > c > 0$ such that $\zeta(s)$ does not vanish on the

domain

$$D = \left\{ s = \sigma + \mathrm{i}\tau \in \mathbb{C} \,\middle|\, \sigma \geq 1 - \frac{c}{1 + \log^+ |\tau|} \right\} \subset \left\{ s \,\middle|\, \sigma > \frac{1}{2} \right\},$$

see *e.g.* [Ten95, Theorem 3.15]. We shall assume $c < \alpha$ in the following.

Proposition 5.1. *Suppose $w = \mathrm{e}^z$ with $|\Re(z)| < \frac{\log 2}{2\,C}$ (hence, w stays in a circular band). The map $s \mapsto G(s,w)$ is holomorphic on the domain D.*

Proof. Since f is additive, one can write for $\sigma > 1$ the Euler product

$$F(s,w) = \prod_{p \in \mathbb{P}} \left(1 + \frac{w^{g(p,1)}}{p^s} + \frac{w^{g(p,2)}}{p^{2s}} + \cdots \right).$$

Notice that by hypothesis on w and s, $1 + \frac{w^{f(p)}}{p^s}$ is non zero for all $p \in \mathbb{P}$, since

$$\left| \frac{w^{f(p)}}{p^s} \right| \leq \frac{\mathrm{e}^{C|\Re(z)|}}{2^{\Re(s)}} \leq 2^{\frac{1}{2} - \Re(s)} < 1.$$

Therefore, the Euler product can be rewritten as

$$F(s,w) = \widetilde{F}(s,w) \, \prod_{p \in \mathbb{P}} \left(1 + \frac{w^{f(p)}}{p^s} \right),$$

where $\widetilde{F}(s,w)$ is uniformly convergent and holomorphic on the domain $\sigma > \frac{1}{2}$. Similarly,

$$\zeta(s)^w = \prod_{p \in \mathbb{P}} \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots \right)^w = Z(s,w) \, \prod_{p \in \mathbb{P}} \left(1 + \frac{w}{p^s} \right)$$

with $Z(s,w)$ holomorphic on the domain $\sigma > \frac{1}{2}$, and non-vanishing on D. It remains therefore to prove that

$$\widetilde{G}(s,w) = \prod_{p \in \mathbb{P}} \frac{1 + \frac{w^{f(p)}}{p^s}}{1 + \frac{w}{p^s}}$$

is itself holomorphic on the domain D. This is clear, since the logarithm of the general term $\log\left(1 + \frac{w^{f(p)}}{p^s}\right) - \log\left(1 + \frac{w}{p^s}\right)$ behaves as

$$\frac{w}{p^s} \, (\mathrm{e}^{z(f(p)-1)} - 1) \simeq \frac{w}{p^s} \, O\!\left(\frac{z}{p^\alpha}\right) = O\!\left(\frac{1}{p^{\Re(s)+\alpha}}\right),$$

and $\Re(s) + \alpha > \Re(s) + c \geq 1$ on the domain D. ☐

The previous Proposition allows one to apply [Ten95, II.5, Theorem 3], which is a Tauberian theorem for $\sum_{k \leq n} w^{f(k)}$. Set

$$Y(s,w) = \frac{1}{s} \, \{(s-1)\, \zeta(s)\}^w = \sum_{j=0}^{\infty} \frac{\gamma_j(w)}{j!} \, (s-1)^j,$$

which is an analytic function in s for $|s-1| < 1$.

Proposition 5.2. *Denote X_n a random integer uniformly chosen in $[1,n]$. One has the mod-Poisson convergence*

$$\mathbb{E}[\mathrm{e}^{zf(X_n)}] = \frac{1}{n} \sum_{k=1}^{n} w^{f(k)} = \mathrm{e}^{(\mathrm{e}^z - 1)\log\log n} \left(\sum_{k=0}^{v} \frac{\lambda_k(\mathrm{e}^z)}{(\log n)^k} + o\!\left(\frac{1}{(\log n)^k}\right) \right),$$

where the coefficients $\lambda_k(w)$ are defined by

$$\lambda_k(w) = \frac{1}{\Gamma(w-k)} \sum_{h+j=k} \frac{1}{h!\,j!} \frac{\partial^h G(1,w)}{\partial s^h} \gamma_j(w).$$

In particular, $\lambda_0(w) = \lim_{s\to 1} \frac{F(s,w)}{\Gamma(w)(\zeta(s))^w}$.

Corollary 5.3. *For* $|\varepsilon| < \frac{\log 2}{2\,C}$*, one has the asymptotics*

$$\mathbb{P}[f(X_n) \geq (1+\varepsilon)\log\log n] \simeq \frac{\mathrm{e}^{-(\log\log n)((1+\varepsilon)\log(1+\varepsilon)-\varepsilon)}}{\sqrt{2\pi \log\log n}} \frac{\sqrt{1+\varepsilon}}{\varepsilon} \lambda_0(\varepsilon).$$

The same estimates have essentially been obtained by Radziwill in [Rad09]. On the other hand, our reasoning makes clear the link between the notion of mod-Poisson convergence, which is a statement on the ratio

$$\frac{\mathbb{E}[\mathrm{e}^{zX_n}]}{(\mathbb{E}[\mathrm{e}^{z\mathcal{P}}])^{\lambda_n}},$$

and the Selberg-Delange method, which is a statement on the ratio

$$\frac{F(s,\mathrm{e}^z)}{(\zeta(s))^{\mathrm{e}^z}}$$

with F the Dirichlet series of an additive arithmetic function.

5.4. **Statistics of random combinatorial objects and singularity analysis.** We now present the method of singularity analysis of generating series which is in spirit comparable to the Selberg-Delange method introduced above but which is used to cover a variety of situations which do not fall in the domain of applications of the Selberg-Delange method. The idea is to use contour integration techniques in order to extract the asymptotic behavior of the coefficients of some generating series of interest. As in the Selberg-Delange method described in the previous section, we shall have two complex variables and the coefficients of our generating series can be themselves the moment generating functions of some random variables. Rather than stating general and abstract results, we would rather illustrate the power of this method (and at the same time the fact that mod-Poisson convergence is genuinely a higher order central limit theorem) by considering the example of the total number of cycles of random permutations under the so called weighted probability measure. For more details the reader can look at the monograph [FS09] or the paper [FO90].

From now on we shall follow the presentation in [NZ13]. Denote $X_n(\sigma)$ the number of disjoint cycles (including fixed points) of a permutation σ in the symmetric group $\mathfrak{S}_n$. We write $C_j(\sigma)$ for the number of cycles of length j in the decomposition of σ as a product of disjoint cycles. Let $\Theta = (\theta_m)_{m\geq 1}$ be given with $\theta_m \geq 0$. The generalized weighted measure is defined as the probability measure $\mathbb{P}_\Theta$ on the symmetric group $\mathfrak{S}_n$:

$$\mathbb{P}_\Theta[\sigma] = \frac{1}{h_n n!} \prod_{m=1}^{n} \theta_m^{C_m(\sigma)}$$

with h_n a normalization constant (or a partition function) and $h_0 = 1$. This model is coming from statistical mechanics and the study of Bose quantum gases (see [NZ13] for more references and details). It generalizes the classical cases of the uniform measure (corresponding to $\theta_m \equiv 1$) and the Ewens measure (corresponding to the case $\theta_m \equiv \theta > 0$). It has been an open question to prove a central limit theorem (as in the Ewens measure

case) for the total number of cycles X_n under such measures (or more precisely under some specific regimes related to the asymptotic behavior of the θ_m's). The difficulty is arising from the fact the there is nothing such as the Feller coupling anymore and that the probabilistic methods fail to hold here. We now show how the method of singularity analysis allows us to prove mod-Poisson convergence, and hence the central limit theorem, but also distributional approximations and precise large deviations. We first consider the generating series

$$g_\Theta(t) = \sum_{n=1}^{\infty} \frac{\theta_n}{n}\, t^n.$$

It is well known that

$$\sum_{n=1}^{\infty} h_n\, t^n = \exp(g_\Theta(t)).$$

Note that in general h_n is not known. Our goal is to obtain an asymptotic for h_n and for the moment generating function of X_n. We note r the radius of convergence of $g_\Theta(t)$. The idea of singularity analysis is to introduce a properly chosen holomorphic function $S(t, w)$ in a domain containing $\{(t, w) \in \mathbb{C}^2;\quad |t| \leq r, |w| \leq \hat{r}\}$, where $\hat{r} > 0$ is some positive number, and then to consider the function $F(t, w) = \exp(wg(t))S(t, w)$ with the goal of extracting precise asymptotic information (with an error term) for the coefficient of t^n in the series expansion as powers of t for $F(t, w)$. This can be carried out if one makes suitable assumptions on the analyticity properties of g together with assumptions on the nature of its singularity at the point r on the circle of convergence. This motivates the next definition:

Definition 5.4. *Let $0 < r < R$ and $0 < \phi < \pi/2$ be given. We then define*

$$\Delta_0 = \Delta_0(r, R, \phi) = \{z \in \mathbb{C};\ |z| < R, z \neq r, |\arg(z - r)| > \phi\}.$$

Assume we are further given $g(t)$, $\theta \geq 0$ and $r > 0$. We then say that $g(t)$ is in the class $\mathcal{F}(r, \theta)$ if

(i) there exists $R > r$ and $0 < \phi < \pi/2$ such that $g(t)$ is holomorphic in $\Delta_0(r, R, \phi)$;

(ii) there exists a constant K such that

$$g(t) = \theta \log\left(\frac{1}{1 - t/r}\right) + K + O(t - r) \quad \text{as } t \to r.$$

One readily notes that the generating series corresponding to the Ewens measure (*i.e.* $\theta_m \equiv \theta$) is of class $\mathcal{F}(1, \theta)$ since in this case

$$g_\Theta(t) \equiv \theta \log\left(\frac{1}{1-t}\right).$$

Consequently our results will provide alternative proofs to this case as well. The next theorem due to Hwang plays a key role in our example (we use the following notation: if $G(t) = \sum_{n=0}^{\infty} g_n t^n$, we denote $[t^n][G] \equiv g_n$ the coefficient of t^n in $G(t)$).

Theorem 5.5 (Hwang, [Hwa94])**.** *Let $F(t, w) = \exp(wg(t))S(t, w)$ be given. Suppose that*

(i) $g(t)$ is of class $\mathcal{F}(r, \theta)$,

(ii) $S(t, w)$ is holomorphic in a domain containing $\{(t, w) \in \mathbb{C}^2;\quad |t| \leq r, |w| \leq \hat{r}\}$, where $\hat{r} > 0$ is some positive number.

Then

$$[t^n][F(t,w)] = \frac{e^{Kw} n^{w\theta-1}}{r^n} \left(\frac{S(r,w)}{\Gamma(\theta w)} + O\left(\frac{1}{n}\right) \right)$$

uniformly for $|w| \leq \hat{r}$ *and with the same* K *as in the definition above.*

The idea of the proof consists in taking a suitable Hankel contour and to estimate the integral over each piece. There exist several other versions of this theorem where one can replace $\log(1 - t/r)$ by other functions and we refer the reader to the monograph [FS09], chapter VI.5. As an application of Theorem 5.5, we obtain an asymptotic for h_n.

Lemma 5.6. *Let* $\theta > 0$ *and assume that* $g_\Theta(t)$ *is of class* $\mathcal{F}(r,\theta)$*. We then have*

$$h_n = \frac{e^K n^{\theta-1}}{r^n \Gamma(\theta)} \left(1 + O\left(\frac{1}{n}\right) \right).$$

Proof. We have already noted that $\sum_{n=1}^{\infty} h_n t^n = \exp(g_\Theta(t))$. We can apply Theorem 5.5 with $g(t) = g_\Theta(t)$, $w = 1$ and $S(t,w) = 1$. □

Now using elementary combinatorial arguments (which are detailed in [NZ13]), one can prove for each $w \in \mathbb{C}$ the following identity as formal power series:

$$\sum_{n=0}^{\infty} h_n \, \mathbb{E}_\Theta[\exp(wX_n)] \, t^n = \exp(e^w g_\Theta(t)). \tag{15}$$

Using this identity and Theorem 5.5, we can show:

Theorem 5.7 (Nikeghbali-Zeindler, [NZ13])**.** *If* $g_\Theta(t)$ *is of class* $\mathcal{F}(r,\theta)$*, then*

$$\mathbb{E}_\Theta[\exp(wX_n)] = n^{\theta(e^w-1)} e^{K(e^w-1)} \left(\frac{\Gamma(\theta)}{\Gamma(\theta e^w)} + O\left(\frac{1}{n}\right) \right).$$

Consequently the sequence (X_n) *converges in the mod-Poisson sense with parameters* $K + \theta \log n$ *and limiting function* $\frac{\Gamma(\theta)}{\Gamma(\theta e^w)}$.

Proof. An application of Theorem 5.5 yields

$$[t^n][\exp(e^w g_\Theta(t))] = \frac{e^{Ke^w} n^{e^w\theta-1}}{r^n} \left(\frac{1}{\Gamma(\theta e^w)} + O\left(\frac{1}{n}\right) \right),$$

with $O(\cdot)$ uniform for bounded $w \in \mathbb{C}$. Now a combination of identity (15) and Lemma 5.6 gives the desired result. □

The above theorem not only implies the central limit theorem, but also Poisson approximations and precise large deviations. We only state here the precise large deviations result which extends earlier work of Hwang in the case $\theta = 1$ as a consequence of Theorem 2.4 and refer to [NZ13] for the distributional approximations results.

Proposition 5.8 (Nikeghbali-Zeindler, [NZ13])**.** *Let* $Y_n = X_n - 1$ *and let* $x \in \mathbb{R}$ *such that* $t_n x \in \mathbb{N}$ *with* $t_n = K + \theta \log n$*. We note* $k = t_n x$*. Then*

$$\mathbb{P}[Y_n = xt_n] = e^{-t_n} \frac{t_n^k}{k!} \left(\frac{\Gamma(\theta)}{x\Gamma(\theta x)} + O\left(\frac{1}{t_n}\right) \right).$$

In fact an application of Theorem 2.4 would immediately yield an arbitrary long expansion for $\mathbb{P}[Y_n = xt_n]$ and also for $\mathbb{P}[Y_n \geq xt_n]$ since the speed of convergence is fast enough.

5.5. **Characteristic polynomials of random matrices in a compact Lie group.** Introduce the classical compact Lie groups of type A, C, D:

$$\begin{aligned}
\mathrm{U}(n) &= \{g \in \mathrm{GL}(n,\mathbb{C}) \mid gg^\dagger = g^\dagger g = I_n\} && \text{(unitary group)}\\
\mathrm{USp}(n) &= \{g \in \mathrm{GL}(n,\mathbb{H}) \mid gg^\star = g^\star g = I_n\} && \text{(compact symplectic group)}\\
\mathrm{SO}(2n) &= \{g \in \mathrm{GL}(2n,\mathbb{R}) \mid gg^t = g^t g = I_{2n}\ ;\ \det g = 1\} && \text{(special orthogonal group)}
\end{aligned}$$

where for compact symplectic groups $g^\star$ denotes the transpose conjugate of a quaternionic matrix, the conjugate of a quaternionic number $a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d$ being $a - \mathrm{i}b - \mathrm{j}b - \mathrm{k}d$. In the following we shall consider quaternionic matrices as complex matrices of size $2n \times 2n$ by using the map

$$a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d \mapsto \begin{pmatrix} a + \mathrm{i}b & c + \mathrm{i}d \\ -c + \mathrm{i}d & a - \mathrm{i}b \end{pmatrix}.$$

The eigenvalues of a matrix $g \in G = \mathrm{SO}(2n)$ or $\mathrm{U}(n)$ or USp(n) are on the unit circle $\mathbb{S}^1$, and for $g \in G$ random taken under the Haar measure

$$\begin{aligned}
\log\det(\mathrm{e}^{\mathrm{i}\theta} - g) &=_{(\mathrm{law})} \log\det \mathrm{e}^{\mathrm{i}\theta} + \log\det(1-g)\\
&= \begin{cases} n\mathrm{i}\theta + \log\det(1-g) & \text{if } G = \mathrm{U}(n),\\ 2n\mathrm{i}\theta + \log\det(1-g) & \text{if } G = \mathrm{SO}(2n) \text{ or } \mathrm{USp}(n), \end{cases}
\end{aligned}$$

so the study of the random characteristic polynomials $\det(z - g)$ reduces to the study of

$$Y_n^{\mathrm{A,C,D}} = \log\det(1-g),$$

which is complex-valued in type A and real-valued in type C, D. In the unitary case, we shall identify $\mathbb{C}$ and $\mathbb{R}^2$ in order to be able to use the multi-dimensional framework introduced in §4. The mean of Y_n is

$$\mathbb{E}[Y_n] = \begin{cases} 0 & \text{for } \mathrm{U}(n);\\ \frac{1}{2}\log\frac{\pi n}{2} & \text{for } \mathrm{USp}(n);\\ \frac{1}{2}\log\frac{8\pi}{n} & \text{for } \mathrm{SO}(2n), \end{cases}$$

and it is proved in [KN12, DKN11] (after [KS00a, KS00b, HKO01]) that $X_n = Y_n - \mathbb{E}[Y_n]$ converges in the mod-Gaussian sense with parameters

$$t_n = \begin{cases} \frac{\log n}{2} & \text{for } \mathrm{U}(n)\\ \log\frac{n}{2} & \text{for } \mathrm{USp}(n) \text{ and } \mathrm{SO}(2n) \end{cases}$$

and

$$\psi = \begin{cases} \frac{G(1+(x+\mathrm{i}y)/2)\,G(1+(x-\mathrm{i}y)/2)}{G(1+x)} & \text{for } \mathrm{U}(n)\\ \frac{G(3/2)}{G(3/2+x)} & \text{for } \mathrm{USp}(n)\\ \frac{G(1/2)}{G(1/2+x)} & \text{for } \mathrm{SO}(2n). \end{cases}$$

Here G denotes Barnes' G-function, which is the entire solution of the functional equation $G(z+1) = \Gamma(z)\,G(z)$ with $G(1) = 1$. Thus, our large deviations theorems apply and one obtains:

Theorem 5.9. *Fix $x > 0$. Over* USp(n)*,*

$$\mathbb{P}_n\left[\det(1-g) \geq \sqrt{\frac{\pi}{2}}\, n^{\frac{1}{2}+x}\right] \simeq \frac{G(3/2)}{x\,G(3/2+x)\,\sqrt{2\pi\log n}}\ \exp\left(-\frac{x^2\,\log 2n}{2}\right);$$

$$\mathbb{P}_n\left[\det(1-g) \leq \sqrt{\frac{\pi}{2}}\, n^{\frac{1}{2}-x}\right] \simeq \frac{G(3/2)}{x\,G(3/2-x)\,\sqrt{2\pi\log n}}\ \exp\left(-\frac{x^2\,\log 2n}{2}\right).$$

Over $\mathrm{SO}(2n)$*,*

$$\mathbb{P}_n\left[\det(1-g) \geq \sqrt{8\pi}\, n^{-\frac{1}{2}+x}\right] \simeq \frac{G(1/2)}{x\, G(1/2+x)\, \sqrt{2\pi \log n}}\, \exp\left(-\frac{x^2\, \log 2n}{2}\right);$$
$$\mathbb{P}_n\left[\det(1-g) \leq \sqrt{8\pi}\, n^{-\frac{1}{2}-x}\right] \simeq \frac{G(1/2)}{x\, G(1/2-x)\, \sqrt{2\pi \log n}}\, \exp\left(-\frac{x^2\, \log 2n}{2}\right).$$

Finally, over $\mathrm{U}(n)$*,*

$$\mathbb{P}_n[|\log\det(1-g)| \geq x\log n] \simeq \left(\frac{1}{2\pi}\int_{\theta=0}^{2\pi} \frac{G(1+x\mathrm{e}^{\mathrm{i}\theta})\, G(1+x\mathrm{e}^{-\mathrm{i}\theta})}{G(1+2x\cos\theta)}\, d\theta\right) \exp(-x^2\log n).$$

Proof. These are immediate computations by using Theorem 3.2 in type C and D, and Theorem 4.12 in type A. In this last case, we also get

$$\begin{aligned}\mathbb{P}_n[\,|\det(1-g)| \geq n^x] &= \mathbb{P}_n[\Re(\log\det(1-g)) \geq x\log n] \\ &\simeq \frac{G(1+x)^2}{G(1+2x)}\, \frac{1}{\sqrt{4\pi\log n}\, x}\, \exp(-x^2\log n)\end{aligned}$$

by restriction to the first variable. □

The analogue of Theorem 5.9 in the setting of random matrices in the β-ensembles, or of general Wigner matrices, has been studied in the recent paper [DE13]. They can be easily restated in the mod-Gaussian language, since their proofs relie on the computation of the asymptotics of the cumulants of the random variables $X_n = \log|\det M_n|$, with for instance $(M_n)_{n\in\mathbb{N}}$ random matrices of the gaussian unitary ensembles.

6. Dependency graphs and mod-Gaussian convergence

Dependency graphs are a classical tool in the literature to prove convergence in distribution towards a Gaussian law of the sum of *partly* dependent random variables. They are used in various domains, such as random graphs [JŁR00, pages 147-152], random polytopes [IV07], patterns in random permutations [Bón10]. As dependency graphs give a natural framework dealing in a uniform way with different kinds of objects, a natural question is the following: when we have a dependency graph with good properties, can we obtain more precise or other kinds of results than the convergence in distribution? Here is a brief presentation of the literature around this question.

- In [BR89], P. Baldi and Y. Rinott give precise estimates for the total variation distance between the relevant sequence of random variables and the Gaussian distribution.
- In [Jan04], S. Janson has established some large deviation result involving the fractional chromatic number of the dependency graph.
- More recently, H. Döring and P. Eichelsbacher have shown how dependency graphs can be used to obtain some moderate deviation principles [DE12, Section 2].

Here, we shall see a link between dependency graphs and mod-Gaussian convergence. This gives us a large collection of examples, for which the material of this article gives automatically some *precise* moderate deviation results. Our deviation result has a larger domain of validity than the one of Döring and Eichelsbacher — see below.

In this section, we establish a general result involving dependency graphs (Theorem 6.3). In the next two sections, we focus on examples and derive the mod-Gaussian convergence of the following renormalized statistics:

- subgraph count statistics in Erdős-Rényi random graphs (Section 7);
- random character values from central measures on partitions (Section 8).

6.1. **The theory of dependency graphs.** Let us consider a variable X, which writes as a sum
$$X = \sum_{\alpha \in V} Y_\alpha$$
of random variables Y_α indexed by a set V.

Definition 6.1. *A graph G with vertex set V is called a* dependency graph *for the family of random variables $\{Y_\alpha,\ \alpha \in V\}$ if the following property is satisfied:*

> *If V_1 and V_2 are disjoint subset of V such that there are no edges in G with one extremity in V_1 and one in V_2, then the sets of random variables $\{Y_\alpha\}_{\alpha \in V_1}$ and $\{Y_\alpha\}_{\alpha \in V_2}$ are independent (*i.e.*, the σ-algebras generated by these sets are independent).*

Note that a family of random variables may admit several dependency graphs. In particular, the complete graph with vertex V is always a dependency graph. However, this is not interesting: the sparser is the dependency graph, the better are the resulting bound. From a dependency graph, one can derive bounds on the cumulants of X; see [Jan88, Lemma 4]

Theorem 6.2. *For any integer $r \geq 1$, there exists a constant C_r with the following property. Let $\{Y_\alpha\}_{\alpha \in V}$ be a family of random variables with dependency graph G. We denote $N = |V|$ the number of vertices of G and D the maximal degree of G. Assume that the variables Y_α have all finite moments and that there exists a constant A such that, for all $\alpha \in V$,*
$$\|Y_\alpha\|_r = \left(\mathbb{E}[(Y_\alpha)^r]\right)^{1/r} \leq A.$$
Then, if $X = \sum_\alpha Y_\alpha$, one has:
$$\kappa^{(r)}(X) \leq C_r\, N\, (D+1)^{r-1}\, A^r.$$

This theorem is often used to prove some central limit theorem. In [DE12], Döring and Eichelsbacher have strengthen this theorem, showing that $C_r = (2\mathrm{e})^r (r!)^3$. Then they have used this new bound to obtain some moderate deviation results. Here, using combinatorial arguments, we show the following:

Theorem 6.3. *Theorem 6.2 holds with $C_r = 2^{r-1}\, r^{r-2}$.*

We shall see in next Section that this stronger version can be used to establish mod-Gaussian convergence and, thus, precise moderate deviation results.

6.2. **Joint cumulants.** There is a multivariate version of cumulants, called joint cumulants, that we shall use to prove Theorem 6.3. We present in this paragraph its definition and basic properties. Most of this material can be found in Leonov's and Shiryaev's paper [LS59] (see also [JŁR00, Proposition 6.16]).

6.2.1. *Preliminaries: set-partitions.* We denote by $[n]$ the set $\{1,\dots,n\}$. A *set partition* of $[n]$ is a (non-ordered) family of non-empty disjoint subsets of S (called parts of the partition), whose union is $[n]$. For instance,

$$\{\{1,3,8\},\{4,6,7\},\{2,5\}\}$$

is a set partition of $[8]$. Denote $\mathfrak{Q}(n)$ the set of set partitions of $[n]$. Then $\mathfrak{Q}(n)$ may be endowed with a natural partial order: the *refinement* order. We say that π is *finer* than π' or π' *coarser* than π (and denote $\pi \leq \pi'$) if every part of π is included in a part of π'.

Lastly, denote μ the Möbius function of the poset $\mathfrak{Q}(n)$. In this paper, we only use evaluations of μ at pairs $(\pi, \{[n]\})$ (the second argument is the partition of $[n]$ in only one part, which is the maximum element of $\mathfrak{Q}(n)$), so we shall use abusively the notation $\mu(\pi)$ for $\mu(\pi,\{[n]\})$. In this case, the value of the Möbius function is given by:

$$\mu(\pi) = \mu(\pi,\{[n]\}) = (-1)^{\#(\pi)-1}(\#(\pi)-1)!. \tag{16}$$

6.2.2. *Definition and properties of joint cumulants.* If $X_1,\dots,X_r$ are random variables with finite moments on the same probability space (denote $\mathbb{E}$ the expectation on this space), we define their joint cumulant by

$$\kappa(X_1,\dots,X_r) = [t_1\cdots t_r]\ \log\left(\mathbb{E}\left[\mathrm{e}^{t_1X_1+\cdots+t_rX_r}\right]\right). \tag{17}$$

As usual, $[t_1\dots t_r]F$ stands for the coefficient of $t_1\cdots t_r$ in the series expansion of F in positive powers of $t_1,\dots,t_r$. Note that joint cumulants are multilinear functions. In the case where all the X_i's are equal, we recover the r-th cumulant $\kappa^{(r)}(X)$ of a single variable, see Section 1. Joint cumulants can be expressed in terms of joint moments, and *vice-versa*:

$$\mathbb{E}[X_1\cdots X_r] = \sum_{\pi\in\mathfrak{Q}(r)}\prod_{C\in\pi}\kappa(X_i\,;\, i\in C); \tag{18}$$

$$\kappa(X_1,\dots,X_r) = \sum_{\pi\in\mathfrak{Q}(r)}\mu(\pi)\prod_{C\in\pi}\mathbb{E}\left[\prod_{i\in C}X_i\right]. \tag{19}$$

In these equations, $C\in\pi$ shall be understood as "C is a part of the set partition π". Recall that $\mu(\pi)$ has an explicit expression given by Equation (16). For example the joint cumulants of one or two variables are simply the mean of a single random variable and the covariance of a couple of random variables:

$$\kappa(X_1) = \mathbb{E}[X_1] \qquad ; \qquad \kappa(X_1,X_2) = \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]\,\mathbb{E}[X_2].$$

For three variables, one has

$$\begin{aligned}\kappa(X_1,X_2,X_3) = \mathbb{E}[X_1X_2X_3] &- \mathbb{E}[X_1X_2]\,\mathbb{E}[X_3] - \mathbb{E}[X_1X_3]\,\mathbb{E}[X_2]\\ &- \mathbb{E}[X_2X_3]\,\mathbb{E}[X_1] + 2\,\mathbb{E}[X_1]\,\mathbb{E}[X_2]\,\mathbb{E}[X_3].\end{aligned}$$

6.2.3. *Statement with joint cumulants.* Let $\{Y_\alpha\}_{\alpha\in V}$ be a family of random variables with dependency graph G. As in Theorem 6.2, we assume that the variables Y_α have all finite moments and that there exists a constant A such that, for all $\alpha\in V$,

$$\|Y_\alpha\|_r = \left(\mathbb{E}[(Y_\alpha)^r]\right)^{1/r} \leq A.$$

Consider r subsets V_1, V_2,... V_r of V, non necessarily distinct and set $X_i = \sum_{\alpha\in V_i} Y_\alpha$ (for $i\in[r]$). We denote D_i the maximal number of vertices in V_i adjacent to a given vertex (not necessarily in V_i). Then one has the following result.

Theorem 6.4. *With the notation above,*

$$|\kappa(X_1, \dots, X_r)| \leq 2^{r-1}\, r^{r-2}\, |V_1|\, (D_2+1) \cdots (D_r+1)\, A^r.$$

The proof of this theorem is very similar to the one of Theorem 6.3. However, to simplify notations, we only prove the latter here.

6.3. **Useful combinatorial lemmas.** We start our proof of Theorem 6.3 by stating a few lemmas on graphs and spanning trees.

6.3.1. *A functional on graphs.* In this Section, we consider graphs H with multiple edges and loops. We use the standard notations $V(H)$ and $E(H)$ for their vertex and edge sets. For a graph H and a set partition π of $V(H)$, we denote $\pi \perp H$ when the following is true: for any edge $\{i,j\} \in E(H)$, the elements i and j lie in different parts of π (notation: $i \nsim_\pi j$). We introduce the following functional on graphs H:

$$\mathrm{F}_H = (-1)^{|V(H)|-1} \sum_{\pi \perp H} \mu(\pi).$$

Lemma 6.5. *For any graph H, one has*

$$\mathrm{F}_H = \sum_{\substack{E \subset E(H) \\ (V(H),E) \text{ connected}}} (-1)^{|E|-|V(H)|+1}.$$

Proof. To simplify notations, suppose $V(H) = [r]$. We denote $\mathbb{1}_{(P)}$ the characteristic function of the property (P). By inclusion-exclusion,

$$\begin{aligned}
(-1)^{|V(H)|-1}\,\mathrm{F}_H &= \sum_{\pi \in \mathfrak{Q}(r)} \left(\prod_{(i,j) \in E(H)} \mathbb{1}_{i \nsim_\pi j} \right) \mu(\pi) = \sum_{\pi \in \mathfrak{Q}(r)} \left(\prod_{(i,j) \in E(H)} (1 - \mathbb{1}_{i \sim_\pi j}) \right) \mu(\pi) \\
&= \sum_{E \subset E(H)} \sum_{\pi \in \mathfrak{Q}(r)} (-1)^{|E|} \left(\prod_{(i,j) \in E} \mathbb{1}_{i \sim_\pi j} \right) \mu(\pi) \\
&= \sum_{E \subset E(H)} (-1)^{|E|} \left[\sum_{\substack{\pi \text{ such that} \\ \forall (i,j) \in E,\ i \sim_\pi j}} \mu(\pi) \right].
\end{aligned}$$

But the quantity in the bracket is 0 unless the only partition in the sum is the maximal partition $\big\{[r]\big\}$, in which case it is 1. This corresponds to the case where the edges in E form a connected subgraph of H. ☐

Corollary 6.6. *The functional F_H fulfills the deletion-contraction induction,* i.e., *if e is an edge of H, then*

$$\mathrm{F}_H = \mathrm{F}_{H/e} + \mathrm{F}_{H \setminus e},$$

where $H \setminus e$ (respectively H/e) are the graphs obtained from H by deleting (resp. contracting) the edge e.

Proof. The first term corresponds to sets of edges containing e, and the second to those that do not contain e. ☐

This induction determines F_H together with the initial conditions:

$$\begin{cases} \mathrm{F}_{\bullet} = 1, \\ \mathrm{F}_{\;} = \mathrm{F}_{\;} = \cdots = 0, \\ \mathrm{F}_H = 0 \quad \text{if } H \text{ is disconnected.} \end{cases}$$

Corollary 6.7. *For any graph H, the quantity F_H is nonnegative and less or equal than the number ST_H of spanning trees in H.*

Proof. The quantity ST_H fulfills the same induction as F_H with initial conditions:

$$\begin{cases} \mathrm{ST}_{\bullet} = 1, \\ \mathrm{ST}_{\;} = \mathrm{ST}_{\;} = \cdots = 1, \\ \mathrm{ST}_H = 0 \quad \text{if } H \text{ is disconnected.} \end{cases}$$
□

Remark 6.8. Both F_H and ST_H are actually specializations of the bivariate *Tutte polynomial* $T_H(x,y)$ of H (*cf.* [Bol98, Chapter X]):

$$\mathrm{F}_H = T_H(1,0) \qquad ; \qquad \mathrm{ST}_H = T_H(1,1).$$

This explains the deletion-contraction relation. As the bivariate Tutte polynomials has non-negative coefficients, it also explains the inequality $0 \leq \mathrm{F}_H \leq \mathrm{ST}_H$.

6.3.2. *Induced graphs containing spanning trees.* Fix a graph G (typically the dependency graphs of our family of variables). For a list $(v_1, \dots, v_r)$ of r vertices of G, we define the induced graph $G[v_1, \dots, v_r]$ as follows:

- its vertex set is $[r]$;
- there is an edge between i and j if and only if $v_i = v_j$ or v_i and v_j are linked in G.

We will be interested in spanning trees of induced graphs. As the vertex set is $[r]$, these spanning trees may be seen as *Cayley trees*. Recall that a Cayley tree of size r is by definition a tree with vertex set $[r]$ (Cayley trees are neither rooted, nor embedded in the plane, they are only specified by an adjacency matrix). These objects are enumerated by the well-known Cayley formula established by C. Borchardt in [Bor60]: there are exactly r^{r-2} Cayley trees of size r.

Lemma 6.9. *Fix a Cayley tree T of size r and a graph G with N vertices and maximal degree D. The number of lists $(v_1, \dots, v_r)$ of r vertices of G such that T is contained in the induced subgraph $G[v_1, \dots, v_r]$ is bounded from above by*

$$N\,(D+1)^{r-1}.$$

Proof. Lists $(v_1, \dots, v_r)$ as in the lemma are constructed as follows. First choose any vertex v_1 among the N vertices. Then consider a neighbor j of 1 in T. As we require $G[v_1, \dots, v_r]$ to contain T, they must also be neighbor in $G[v_1, \dots, v_r]$, that is to say that $v_j = v_1$ or v_j is a neighbour of v_1 in T. Thus, once v_1 is fixed, there are at most $D+1$ possible values for v_j. The same is true for all neighbors of 1 and then for all neighbors of neighbors of 1 and so on. □

We have the following immediate consequence.

Corollary 6.10. *Let G be a graph with n vertices and maximal degree D and $r \geq 1$. The number of couples $\big((v_1, \ldots, v_r), T\big)$ where each v_i is a vertex of V and T a spanning tree of the induced subgraph $G[v_1, \ldots, v_r]$ is bounded above by*

$$r^{r-2}\, N\, (D+1)^{r-1}.$$

6.3.3. *Spanning trees and set partitions of vertices.* Recall that ST_H denotes the number of spanning trees of a graph H. Consider now a graph H with vertex set $[r]$ and a set partition $\pi = (\pi_1, \ldots, \pi_t)$ of $[r]$. For each i, we denote $\mathrm{ST}^{\pi_i}(H) = \mathrm{ST}_{H[\pi_i]}$ the number of spanning trees of the graph induced by H on the vertex set π_i. We also use the multiplicative notation

$$\mathrm{ST}^{\pi}(H) = \prod_{j=1}^{t} \mathrm{ST}^{\pi_j}(H).$$

We can also look at the contraction H/π of H with respect to π. By definition, it is the multigraph (*i.e.* graph with multiple edges, but no loops) defined as follows. Its vertex set is the index set $[t]$ of the parts of π and, for $i \neq j$, there are as many edges between i and j as edges between a vertex of π_i and a vertex of π_j in H. Denote $\mathrm{ST}_{\pi}(H) = \mathrm{ST}_{H/\pi}$ the number of spanning trees of this contracted graph (multiple edges are here important). This should not be confused with $\mathrm{ST}^{\pi}(H)$: in the latter, π is placed as an exponent because the quantity is multiplicative with respect to the part of π.

Note that the union of a spanning tree $\overline{T}$ of H/π and of spanning trees T_i of $H[\pi_i]$ (for $1 \leq i \leq t$) gives a spanning tree T of H. Conversely, take a spanning tree T on H and a bicoloration of its edges. Edges of color 1 can be seen as a subgraph of H with the same vertex set $[r]$. This graph is of course acyclic. Its connected components define a partition $\pi = \{\pi_1, \ldots, \pi_t\}$ of $[r]$ and edges of color 1 correspond to a collection of spanning trees T_i of $H[\pi_i]$ (for $1 \leq i \leq t$). Besides, edges of color 2 define a spanning tree $\overline{T}$ on H/π.

Therefore, we have described a bijection between spanning trees of H with a bicoloration of their edges and triple $(\pi, \overline{T}, (T_i)_{1 \leq i \leq t})$ where:

- π is a set partition of the vertex set $[r]$ of H (we denote t its number of parts);
- $\overline{T}$ is a spanning tree of the contracted graph H/π;
- for each $1 \leq i \leq t$, T_i is a spanning tree of the induced graph $H[\pi_i]$.

Before giving a detailed example, let us state the numerical corollary of this bijection:

$$2^{r-1}\, \mathrm{ST}_H = \sum_{\pi} \mathrm{ST}_{\pi}(H)\, \mathrm{ST}^{\pi}(H), \tag{20}$$

where the sum runs over all set partitions π of $[r]$.

Our bijection is illustrated on Figure 9, with the following conventions:

- red dotted edges belong to the graph but not to the spanning tree;
- blue plain edges are edges of color 1 in the tree;
- green dashed edges are edges of color 2 in the tree.

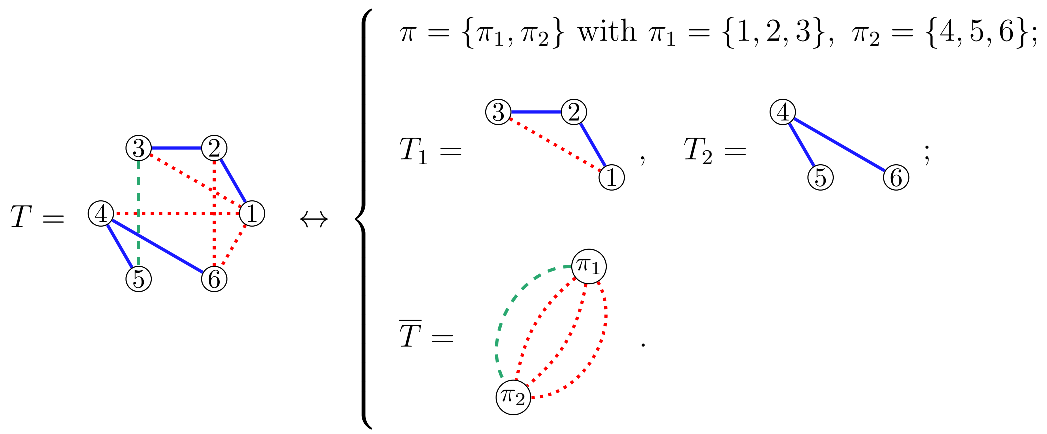


FIGURE 9. Bijection explaining Identity (20).

Note that in this example the graph H/π has two vertices linked by four edges. These four edges correspond to the edges $(1, 4)$, $(1, 6)$, $(2, 6)$ and $(3, 5)$ of H. In the example, the spanning tree $\overline{T}$ is the edge $\{(3, 5)\}$. If we had chosen another edge, the tree T on the left-hand side would have been different. Hence, in the Equality (20), the multiple edges of H/π must be taken into account: in our example, $\mathrm{ST}_\pi(H) = 4$.

6.4. **Proof of the bound on cumulants.** Recall that we want to find a bound for $\kappa^{(r)}(X)$. As X writes $X = \sum_{\alpha \in V} Y_\alpha$, we may use joint cumulants and expand by multilinearity:

$$\kappa^{(r)}(X) = \sum_{\alpha_1, \ldots, \alpha_r} \kappa(Y_{\alpha_1}, \ldots, Y_{\alpha_r}). \tag{21}$$

The sum runs over lists of r elements in V, that is vertices of the dependency graph G. The proof consists in bounding each summand $\kappa(Y_{\alpha_1}, \ldots, Y_{\alpha_r})$, with a bound depending on the induced subgraph $G[\alpha_1, \ldots, \alpha_r]$.

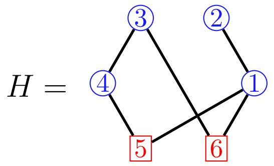
6.4.1. *Bringing terms together in joint cumulants.* Recall the moment-cumulant formula (19) which states $\kappa(Y_{\alpha_1}, \ldots, Y_{\alpha_r}) = \sum_\pi \mu(\pi) M_\pi$, where

$$M_\pi = \prod_{B \in \pi} \mathbb{E}\left[\prod_{i \in B} Y_{\alpha_i}\right].$$

We warn the reader that the notation M_π is a little bit abusive as this quantity depends also on the list $(\alpha_1, \ldots, \alpha_r)$. By hypothesis, G is a dependency graph for the family $\{Y_\alpha\}_{\alpha \in V}$. Hence if some block B of a partition π can be split into two sub-blocks B_1 and B_2 such that the set of vertices $\{\alpha_i\}_{i \in B_1}$ and $\{\alpha_i\}_{i \in B_2}$ are disjoint and do not share an edge, then

$$\mathbb{E}\left[\prod_{i \in B} Y_{\alpha_i}\right] = \mathbb{E}\left[\prod_{i \in B_1} Y_{\alpha_i}\right] \times \mathbb{E}\left[\prod_{i \in B_2} Y_{\alpha_i}\right]. \tag{22}$$

Therefore, $M_\pi = M_{\phi_H(\pi)}$, where $H = G[\alpha_1, \ldots, \alpha_r]$ and $\phi_H(\pi)$ is the refinement of π obtained as follows: for each part π_i of π, consider the induced graph $H[\pi_i]$ and replace π_i by the collection of vertex sets of the connected components of $H[\pi_i]$. This construction is illustrated on Figure 10.



For example, consider the graph H here opposite and the partition $\pi = \{\pi_1, \pi_2\}$ with $\pi_1 = \{1,2,3,4\}$ and $\pi_2 = \{5,6\}$. Then $H[\pi_1]$ (respectively, $H[\pi_2]$) has two connected components with vertex sets $\{1,2\}$ and $\{3,4\}$ (resp. $\{5\}$ and $\{6\}$). Thus

$$\phi_H(\pi) = \{\{1,2\},\{3,4\},\{5\},\{6\}\}.$$

FIGURE 10. Illustration of the definition of ϕ_H.

We can thus write

$$\kappa(G) = \sum_{\pi'} M_{\pi'} \left(\sum_{\pi \in \phi_H^{-1}(\pi')} \mu(\pi) \right).$$

Fix $\pi' = (\pi'_1, \dots, \pi'_t)$ and let us have a closer look to the expression in the parentheses that we will call $\alpha_{\pi'}$. To compute it, it is convenient to consider the contraction H/π' of the graph H with respect to the partition π'.

Lemma 6.11. *Let π' be a set partition of $[r]$. If one of the induced graph $H[\pi'_i]$ is disconnected, then $\alpha_{\pi'} = 0$. Otherwise, $\alpha_{\pi'} = \mathrm{F}_{H/\pi'}$.*

Proof. The first part is immediate, as $\phi_H^{-1}(\pi') = \emptyset$ in this case.

If all induced graphs are connected, let us try to describe $\phi_H^{-1}(\pi')$. All set partitions π of this set are coarser than π', so can be seen as set partitions of the index set $[r]$ of the parts of π'. This identification does not change their Möbius functions, which depends only on the number of parts. Then, it is easy to see that π lies in $\phi_H^{-1}(\pi')$ if and only if π is coarser than π' and two elements in the same part of π never share an edge in H/π' (here, π is seen as a set partition of $[r]$). In other words, π lies in $\phi_H^{-1}(\pi')$ if and only if $\pi \perp (H/\pi')$. This implies the Lemma. □

Consequently,

$$\kappa(Y_{\alpha_1}, \dots, Y_{\alpha_r}) = \sum_{\pi'} M_{\pi'}\, \mathrm{F}_{H/\pi'} \left(\prod_{i=1}^{t} \mathbb{1}_{H[\pi'_i] \text{ connected}} \right), \tag{23}$$

where the sum runs over all set partitions π' of $[r]$.

6.4.2. *Bounding all the relevant quantities.* The following bound for M_π follows directly from Hölder inequality and the assumption $\|Y_\alpha\|_r \le A$ (for every $\alpha \in V$):

$$|M_\pi| \le A^r. \tag{24}$$

Finally, to bound each summand $\kappa(Y_{\alpha_1}, \dots, Y_{\alpha_r})$, we shall use the following bounds:

$$|\mathrm{F}_{H/\pi'}| \le \mathrm{ST}_{H/\pi'} \qquad \text{by Corollary 6.7;}$$
$$\mathbb{1}_{H[\pi'_i] \text{ connected}} \le \mathrm{ST}_{H[\pi'_i]}.$$

Thus, Equation (23) gives

$$|\kappa(Y_{\alpha_1}, \dots, Y_{\alpha_r})| \le A^r \sum_{\pi'} \mathrm{ST}_{H/\pi'} \left(\prod_{i=1}^{t} \mathrm{ST}_{H[\pi'_i]} \right) = A^r\, 2^{r-1}\, \mathrm{ST}_H, \tag{25}$$

the last equality corresponding to Equation (20). Using Equation (21), we get that the cumulant $\kappa^{(r)}(X)$ is smaller than $2^{r-1}A^r$ times the number of couples $((\alpha_1,\dots,\alpha_r),T)$ where $(\alpha_1,\dots,\alpha_r)$ is a list of vertices of the dependency graph G and T a spanning tree of the induced subgraph $G[\alpha_1,\dots,\alpha_r]$.

Corollary 6.10 now ends the proof of Theorem 6.3.

7. Subgraph counts in Erdős-Rényi random graphs

In this section, we consider Erdős-Rényi model $\Gamma(n,p_n)$ of random graphs. A random graph Γ with this distribution is described as follows. Its vertex set is $[n]$ and for each pair $\{i,j\}\subset[n]$ with $i\neq j$, there is an edge between i and j with probability p_n. Moreover, all these events are independent. We are then interested in the following random variables, called *subgraph count statistics*. If γ is a fixed graph of size k, then $X_\gamma^{(n)}$ is the number of copies of γ *contained* in the graph $\Gamma(n,p_n)$ (a more formal definition is given in the next paragraph). This is a classical parameter in random graph theory; see, *e.g.* the book of S. Janson, T. Łuczak and A. Ruciński [JŁR00].

The first result on this parameter was obtained by P. Erdős and A. Rényi, *cf.* [ER60]. They proved that, if γ belongs to some particular family of graphs (called *balanced*), one has a threshold:

$$\lim_{n\to\infty}\mathbb{P}[X_\gamma^{(n)}>0]=\begin{cases}0 & \text{if } p_n=o\left(n^{-1/m(\gamma)}\right);\\ 1 & \text{if } n^{-1/m(\gamma)}=o(p_n),\end{cases}$$

where $m(\gamma)=|E(\gamma)|/|V(\gamma)|$. This result was then generalized to all graphs by B. Bollobás in [Bol01], but the parameter $m(\gamma)$ is in general more complicated than the quotient above. Consider the case $n^{-1/m(\gamma)}=o(p_n)$, when the graph $\Gamma(n,p_n)$ contains with high probability a copy of γ. It was then proved by A. Ruciński (see [Ruc88]) that, under the additional assumption $n^2(1-p_n)\to\infty$, the fluctuations of $X_\gamma^{(n)}$ are Gaussian. This result can be obtained using dependency graphs; see *e.g.* [JŁR00, pages 147-152].

Here, we consider the case where $p_n=p$ is a constant sequence ($0<p<1$). The possibility of relaxing this hypothesis is discussed in Section 7.3.3. Denote $\alpha_n=n^2$ and $\beta_n=n^{k-2}$, where k is the number of vertices of γ. It is easy to check that

$$\mathbb{E}[X_\gamma^{(n)}]=c\,\alpha_n\,\beta_n\quad;\quad \mathrm{Var}(X_\gamma^{(n)})=\sigma^2\,\alpha_n\,(\beta_n)^2$$

for some positive constants c and σ — see, *e.g.*, [JŁR00, Lemma 3.5]. Hence, Ruciński's central limit theorem asserts that, if $T\sim x\sqrt{\alpha_n}$ for some fixed real x, then

$$\lim_{n\to\infty}\mathbb{P}\left[\frac{X_\gamma^{(n)}-\mathbb{E}[X_\gamma^{(n)}]}{\beta_n}\geq T\right]=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x/\sigma}e^{-u^2/2}du.$$

Using Theorem 6.3, we shall extend this result to a framework where x tends to infinity, but not to quickly: $(\alpha_n)^{1/2}\ll T\ll(\alpha_n)^{3/4}$.

Theorem 7.1. *Let $0<p<1$ and γ be a graph with k vertices. We consider $X_\gamma^{(n)}$ the number of copies of γ contained in Erdős-Rényi random graph $\Gamma(n,p)$. Let α_n and β_n be defined as above.*

(1) *The renormalized variable* $(X_\gamma^{(n)} - \mathbb{E}[X_\gamma^{(n)}])/(\alpha_n^{1/3}\beta_n)$ *converges mod-Gaussian with parameters* $t_n = \sigma^2\,\alpha_n^{1/3}$ *and limiting function* $\psi(z) = \exp\left(\frac{L}{6}z^3\right)$*, where* σ *and* L *are computed in §7.3.2.*

(2) *Besides, if* $(\alpha_n)^{1/2} \ll T \ll (\alpha_n)^{3/4}$*, the following precise moderate deviations principle holds:*

$$\mathbb{P}\left[\frac{X_\gamma^{(n)} - \mathbb{E}[X_\gamma^{(n)}]}{\beta_n} \geq T\right] = \frac{\mathrm{e}^{-\frac{T^2}{2\sigma^2\alpha_n}}}{\sqrt{2\pi\sigma^2\frac{T^2}{\alpha_n}}}\,\exp\left(\frac{L\,T^3}{6\,\sigma^3\,(\alpha_n)^2}\right)\big(1+o(1)\big). \tag{26}$$

A similar result has been obtained by H. Döring and P. Eichelsbacher in [DE12, Theorem 2.3]. However,

- their result is less precise as they only obtain the equivalence of the logarithms of the relevant quantities (in particular, when we look at the logarithm, the second factor of the right-hand side is negligible);
- and their proof works for $T \ll n^{6/5}$ while ours is valid for $T \ll n^{3/2}$.

Unfortunately, we cannot get deviation results when $T \sim xn^2$ for some real number x; this would correspond to evaluate $\mathbb{P}[X_\gamma^{(n)} > (1+\varepsilon)\,\mathbb{E}[X_\gamma^{(n)}]]$. For large deviations equivalents of

$$\log\,\mathbb{P}[X_\gamma^{(n)} > (1+\varepsilon)\,\mathbb{E}[X_\gamma^{(n)}]]$$

with $X_\gamma^{(n)}$ the subgraph count in a Erdős-Rényi random graph, there is a quite large literature, see [CV11, Theorem 4.1] and [Cha12] for recent results in this field. As we consider deviations of a different scale, our result is neither implied by, nor implies these results. Note, however, that their large deviation results are equivalents of the logarithm of the probability, while our statement is an equivalent for the probability itself.

7.1. **A bound on cumulants.**

7.1.1. *Subgraph count statistics.* In the following we denote $\mathfrak{A}(n,k)$ the set of *arrangements* in $[n]$ of length k, *i.e.*, lists of k distinct elements in $[n]$. The cardinality of $\mathfrak{A}(n,k)$ is the falling factorial $n^{\downarrow k} = n(n-1)\cdots(n-k+1)$. Let $A = (a_1,\dots,a_k)$ be an arrangement in $[n]$ of length k, and γ be a fixed graph with vertex set $[k]$. Recall that $\Gamma = \Gamma(n,p_n)$ is a random Erdős-Rényi graph on $[n]$. We denote $\delta_\gamma(A)$ the following random variable:

$$\delta_\gamma(A) = \begin{cases} 1 & \text{if } \gamma \subseteq \Gamma[a_1,\dots,a_k]; \\ 0 & \text{else.} \end{cases} \tag{27}$$

Here $\Gamma[a_1,\dots,a_k]$ denotes the graph induced by Γ on vertex set $\{a_1,\dots,a_k\}$. As our data is an *ordered* list $(a_1,\dots,a_k)$, this graph can canonically be seen as a graph on vertex set $[k]$, that is the same vertex set as γ. Then the inclusion should be understood as inclusion of edge sets.

For any graph γ and any integer $n \geq 1$, we then define the random variable $X_\gamma^{(n)}$ by

$$X_\gamma^{(n)} = \sum_{A \in \mathfrak{A}(n,|\gamma|)} \delta_\gamma(A).$$

Remark 7.2. It would also be natural to replace in Definition (27) the inclusion by an equality $\gamma = \Gamma[a_1, \dots, a_k]$. This would lead to other random variables $Y_\gamma^{(n)}$, called *induced* subgraph counts. Their asymptotic behavior is harder to study (in particular, fluctuations are not always Gaussian; see [JŁR00, Theorem 6.52]). Notice however that if γ is a complete graph, then the two definitions correspond. Hence, our results include the numbers of k-cliques in a random Erdős-Rényi graph.

7.1.2. *A dependency graph of the subgraph count statistics.* Fix some graph γ with vertex set $[k]$. By definition, the variable we are interested in writes as a sum

$$X_\gamma^{(n)} = \sum_{A \in \mathfrak{A}(n,k)} \delta_\gamma(A).$$

We shall describe a dependency graph for the variables $\{\delta_\gamma(A)\}_{A \in \mathfrak{A}(n,k)}$.

For each pair $e = \{v, v'\} \subset [n]$, denote I_e the indicator function of the event: *e is in the graph* $\Gamma(n,p)$. By definition of the model $\Gamma(n,p)$, the random variables I_e are independent Bernouilli variables of parameter p. Then, for an arrangement A, denote $E(A)$ the set of pairs $\{v, v'\}$ where v and v' appear in the arrangement A. One has

$$\delta_\gamma(A) = \prod_{e \in E'(A)} I_e,$$

where $E'(A)$ is a subset of $E(A)$ determined by the graph γ. In particular, if, for two arrangements A and A', one has $|E(A) \cap E(A')| = \emptyset$ (equivalently, $|A \cap A'| \leq 1$), then the variables $\delta_\gamma(A)$ and $\delta_\gamma(A')$ are defined using different variables I_e (and, hence, are independent). This implies that the following graph denoted $\mathbb{B}$ is a dependency graph for the family of variables $\{\delta_\gamma(A)\}_{A \in \mathfrak{A}(n,k)}$:

- its vertex set is $\mathfrak{A}(n,k)$;
- there is an edge between A and A' if $|A \cap A'| \geq 2$.

Considering this dependency graph is quite classical – see, *e.g.*, [JŁR00, Example 1.6].

All variables in this graph are Bernouilli variables and, hence, bounded by 1. Besides the graph $\mathbb{B}$ has $N = n^{\downarrow k}$ vertices, and is regular of degree D smaller than

$$\binom{k}{2}^2 2\,(n-2)(n-3)\dots(n-k-1) < k^4\, n^{k-2}.$$

Indeed, a neighbour A' of a fixed arrangement $A \in \mathfrak{A}(n,k)$ is given as follows:

- choose a pair $\{a_i, a_j\}$ in A that will appear in A';
- choose indices i' and j' such that $a'_{i'} = a_i$ and $a'_{j'} = a_j$ (these indices are different but their order matters);
- choose the other values in the arrangement A'.

So, we may apply Theorem 6.3 and we get:

Proposition 7.3. *Fix a graph γ of vertex set $[k]$. For any $r \leq 1$, one has*

$$\left|\kappa^{(r)}(X_\gamma^{(n)})\right| \leq 2^{r-1} r^{r-2}\, n^k\, (k^4\, n^{k-2})^{r-1}.$$

7.2. **Polynomiality of cumulants.**

7.2.1. *Dealing with several arrangements.* Consider a list $(A^1, \dots, A^r)$ of arrangements. We associate to this data two graphs (unless said explicitly, we always consider loopless simple graphs, and $V(G)$ and $E(G)$ denote respectively the edge and vertex sets of a graph G):

- the graph $G_{\mathbb{A}}$ has vertex set
$$V_{\mathbb{k}} = V(G_{\mathbb{A}}) = \{(t,i) \mid 1 \leq i \leq r, 1 \leq t \leq k_i\}$$
and an edge between (t,i) and (s,j) if and only if $a^i_t = a^j_s$. It is always a disjoint union of cliques. If the A^i's are arrangements, then the graph $G_{\mathbb{A}}$ is endowed with a natural proper r-coloring, (t,i) being of color i.
- the graph $H^m_{\mathbb{A}}$ has vertex set $[r]$ and an edge between i and j if $|A^i \cap A^j| \geq m$.

Notice that $H^1_{\mathbb{A}}$ is the *contraction* of the graph $G_{\mathbb{A}}$ by the map $\varphi : (t,i) \mapsto i$ from the vertex set of $G_{\mathbb{A}}$ to the vertex of $H^1_{\mathbb{A}}$. Indeed,
$$(i,j) \in E(H^1_{\mathbb{A}}) \;\Leftrightarrow\; \exists v_i \in \varphi^{-1}(i), v_j \in \varphi^{-1}(j) \text{ such that } (v_i, v_j) \in E(G_{\mathbb{A}}).$$

An example of a graph $G_{\mathbb{A}}$ and of its (1-)contraction $H_{\mathbb{A}}$ is drawn on Figure 11. For $m \geq 2$, the definition of $H^m_{\mathbb{A}}$ is less common. We call it the m-contraction of $G_{\mathbb{A}}$. The 2-contraction is interesting for us. It corresponds exactly to the graph induced by the dependency graph $\mathbb{B}$ on the list of arrangement $\mathbb{A}$, considered in the proof of Theorem 6.3.

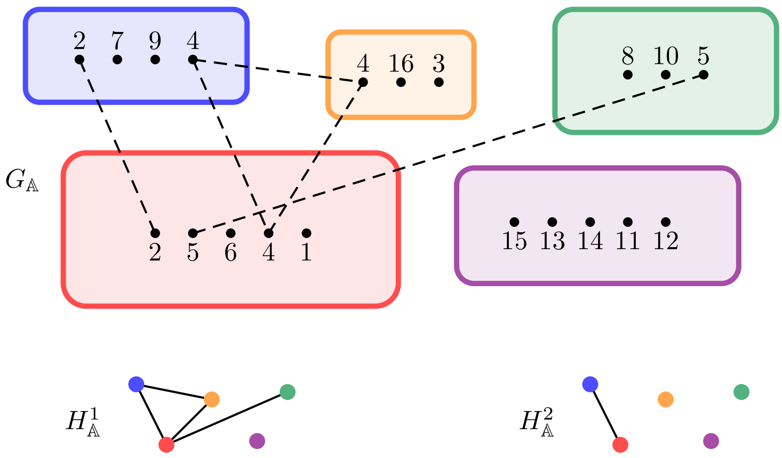


FIGURE 11. The graphs $G_{\mathbb{A}}$, $H^1_{\mathbb{A}}$ and $H^2_{\mathbb{A}}$ corresponding to the family of arrangements $(2,5,6,4,1)$, $(2,7,9,4)$, $(4,16,3)$, $(8,10,5)$ $(15,13,14,11,12)$.

Remark 7.4. Graphs associated to families of arrangements are a practical way to encode some information and should not be confused with the random graphs or their induced subgraphs. Therefore we used greek letters for the latter and latin letter for graphs $G_{\mathbb{A}}$ and their contractions. The dependency graph will always be called $\mathbb{B}$ to avoid confusions.

7.2.2. *Exploiting symmetries.* The dependency graph of our model has much more structure than a general dependency graph. In particular, all variables $\delta_\gamma(A)$ are identically distributed. More generally, the joint distribution of

$$\big(\delta_\gamma(A^1), \dots, \delta_\gamma(A^r)\big)$$

depends only of the 2−contraction of $G_{\mathbb{A}}$. Here, we state a few consequences of this invariance property that will be useful in the next Section.

Lemma 7.5. *Fix a graph γ of size k, the quantity*

$$\mathbb{E}\big[\delta_\gamma(A^1) \cdots \delta_\gamma(A^r)\big]$$

depends only on the graph $G_{\mathbb{A}}$ associated to the family of arrangements $(A^1, \dots, A^r)$. The same is true for the joint cumulant $\kappa\left(\delta_\gamma(A^1), \dots, \delta_\gamma(A^r)\right)$.

Proof. The first statement follows immediately from the invariance of the model $\Gamma(n, p)$ by relabelling of the vertices. The second is a corollary, using the moment-cumulant relation (19). □

Corollary 7.6. *Fix some graph γ. Then the joint cumulant $\kappa(X_\gamma^{(n)}, \dots, X_\gamma^{(n)})$ is a polynomial in n.*

Proof. Using Lemma 7.5, we can rewrite the expansion (21) as

$$\kappa(X_\gamma^{(n)}, \dots, X_\gamma^{(n)}) = \sum_G \kappa(G) N_G, \tag{28}$$

where:

- the sum runs over graphs G of vertex set $V_{\mathbb{k}}$ that correspond to some arrangements (that is G is a disjoint union of cliques and, for any s, t and i, there is no edge between (s, i) and (t, i));
- $\kappa(G)$ is the common value of $\kappa\left(\delta_\gamma(A^1), \dots, \delta_\gamma(A^r)\right)$, where $(A^1, \dots, A^r)$ is any list of arrangements with associated graph G;
- N_G is the number of lists of arrangements with associated graph G.

But it is clear that the sum index is finite and that neither the summation index nor the quantity $\kappa(G)$ depend on n. Besides, the number N_G is simply the falling factorial $n(n-1)\dots(n-c_G+1)$, where c_G is the number of connected components of G. The corollary follows from these observations. □

7.3. **Moderate deviations for subgraph counts.**

7.3.1. *End of the proof of Theorem 7.1.* We would like to apply Corollary 3.6 to the sequence $S_n = X_\gamma^{(n)} - \mathbb{E}[X_\gamma^{(n)}]$ with $\alpha_n = n^2$ and $\beta_n = n^{k-2}$. Let us check that S_n indeed fulfills the hypothesis.

(1) The uniform bound $|\kappa^{(r)}(S_n)| \leq (Cr)^r \alpha_n (\beta_n)^r$, where C does not depend on n, corresponds to Proposition 7.3; we may even choose $C = 2k^4$.

(2) We also have to check the speed of convergence:

$$\kappa^{(2)}(S_n) = \sigma^2\, \alpha_n\, (\beta_n)^2\, (1 + o(\alpha_n^{-5/12})) \quad ; \quad \kappa^{(3)}(S_n) = L\, \alpha_n\, (\beta_n)^3\, (1 + o(\alpha_n^{-1/6})). \qquad (29)$$

But these estimates follow directly from the bound above for $r = 2, 3$ and the fact that $\kappa^{(r)}(S_n)$ is always a polynomial in n — see Corollary 7.6.

Finally, the mod-Gaussian convergence follows from the observations in Section 3.2 and we can apply Corollary 3.6 to get the moderate deviation statement. This ends the proof of Theorem 7.1. □

Remark 7.7. Using Theorem 6.4, we could obtain a bound for joint cumulant of subgraph counts. Hence, it would be possible to derive mod-Gaussian convergence and moderate deviation results for *linear combinations and vectors* of subgraph counts, in the spirit of Section 4. However, as we do not have a specific motivation for that and as the statement for a single subgraph count statistics is already quite technical, we have chosen not to present such a result.

7.3.2. *Computing σ^2 and L.* The proof above does not give an explicit value for σ^2 and L. Yet, these values can be obtained by analyzing the graphs G that contribute to the highest degree term of $\kappa^{(2)}$ and $\kappa^{(3)}$.

Lemma 7.8. *Let γ be a graph with k vertices and h edges. Then the positive number σ appearing in Theorem 7.1 is given by*

$$\sigma^2 = 2h^2 p^{2h-1}(1-p).$$

Proof. By definition, σ^2 is the coefficient of n^{2k-2} in $\kappa^{(2)}(X_\gamma^{(n)})$. As seen in Equation (28), the quantity $\kappa^{(2)}(X_\gamma^{(n)})$ can be written as

$$\sum_G \kappa(G) N_G,$$

where the sum runs over some graphs G with vertex set $V \sqcup V$. However, we have seen that $\kappa(G) = 0$ unless the 2-contraction H^2 of G is connected — see Inequality (25) — and on the other hand, N_G is a polynomial in n, whose degree is the number c_G of connected component of G.

As we are interested in the coefficient of n^{2k-2}, we should consider only graphs G with at least $2k-2$ connected components and a connected 2-contraction. These graphs are represented on Figure 12.

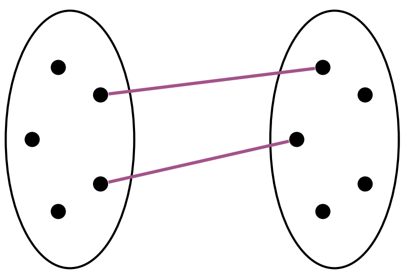


FIGURE 12. Graphs involved in the computation of the main term in $\kappa^{(2)}(X_\gamma^{(n)})$.

Namely, we have to choose a pair of vertices on each side and connect each of these vertices to one vertex of the other pair (there are 2 ways to make this connection, if both pairs are fixed). A quick computation shows that, for such a graph G,

$$\kappa(G) = \begin{cases} p^{2h-1}(1-p) & \text{if both pairs correspond to an edge of } \gamma; \\ 0 & \text{else.} \end{cases}$$

Finally there are $2h^2$ graphs with a non-zero contribution to the coefficient of n^{2k-2} in $\kappa^{(2)}(X_\gamma^{(n)})$. For each of these graph, $N_G = n(n-1)\dots(n-2k+3) = n^{2k-2}(1+o(1))$. Therefore, the coefficient of n^{2k-2} in $\kappa^{(2)}(X_\gamma^{(n)})$ is $2h^2p^{2h-1}(1-p)$, as claimed. □

The number L can be computed by the same method.

Lemma 7.9. *Let γ be a graph with k vertices and h edges. Then the number L appearing in Theorem 7.1 is given by*

$$L = 12h^3(h-1)p^{3h-2}(1-p)^2 + 4h^3p^{3h-2}(1-p)(1-2p).$$

Proof. Here, we have to consider graphs G on vertex set $V \sqcup V \sqcup V$ with at least $3k-4$ connected components and with a connected 2-contraction. These graphs are of two kinds, see Figure 13.

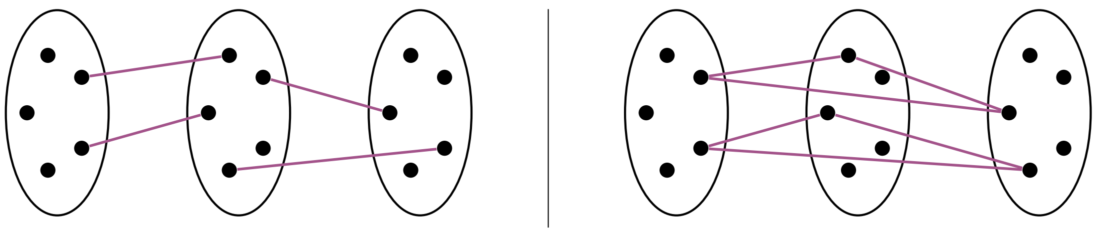


FIGURE 13. Graphs involved in the computation of the main term in $\kappa^{(3)}(X_\gamma^{(n)})$.

In the first case (left-hand side picture), an edge on the left can possibly have an extremity in common with an edge on the right. In this case, one has to add an edge to complete the triangle (indeed, all graphs G are disjoint unions of cliques). But this cannot happen for both edges on the left simultaneously, otherwise the graph belong to the second family.

The following is now easy to check. There are $12h^3(h-1)$ graphs of the first kind with a non-zero cumulant $\kappa(G)$ — 3 choices for which set V plays the central role, $h^3(h-1)$ for pairs of vertices and 4 ways to link the chosen vertices — and, for these graphs, the corresponding cumulant is always $\kappa(G) = p^{3h-2}(1-p)^2$. Similarly, there are $4h^3$ graphs of the second kind with a non-zero cumulant $\kappa(G)$. For these graphs, $\kappa(G) = p^{3h-2}(1-p)(1-2p)$. In both cases, $N_G = n^{3k-4}(1+o(1))$. This completes the proof. □

Example 7.10. Denote T_n the number of triangles in a random Erdős-Rényi graph $\Gamma(n,p)$. According to the previous Lemmas, the parameters σ^2 and L are respectively

$$\sigma^2 = 18\,p^5\,(1-p) \quad \text{and} \quad L = 108\,p^7\,(1-p)(7-8p).$$

Moreover, $\mathbb{E}[T_n] = n^{\downarrow 3}\, p^3 = n^3\, p^3 - 3\, n^2\, p^3 + O(n)$. So,

$$\mathbb{P}\big[T_n \geq n^3 p^3 + n^2(v - 3p^3)\big] \simeq \frac{1}{\sqrt{36\pi\, p^5 (1-p) v^2}} \exp\left(-\frac{v^2}{36\, p^5\, (1-p)} + \frac{(7-8p)\, v^3}{3n\sqrt{2p(1-p)}} \right)$$

for $1 \ll v \ll n^{1/2}$.

7.3.3. *Case of non-constant sequences p_n.* Proposition 7.3 still holds when p_n is a non-constant sequence (the particularly interesting case is $p_n \to 0$). One can even sharpen a little this bound by replacing Inequality (24) by the trivial bound

$$|M_\pi| \leq (p_n)^h,$$

where h is the number of edges of γ. Doing that, we get

$$\left|\kappa^{(r)}(X_\gamma^{(n)})\right| \leq C^r\, r^{r-2}\, n^{r(k-2)+2}\, (p_n)^h.$$

But, unlike in the case $p_n = p$ constant, this bound is not always optimal (up to a multiplicative constant) for a fixed r. For example, if $p_n = n^{-\epsilon}$ with $\epsilon > 0$ sufficiently small, then — see [JŁR00, Lemma 3.5] —

$$\mathrm{Var}(X_\gamma^{(n)}) \sim \text{const.} \times n^{2k-2} (p_n)^{2h-1} \ll n^{2k-2} (p_n)^h.$$

Finding a uniform bound for cumulants, whose dependence in r is of order $(Cr)^r$ (so that we have mod-Gaussian convergence), and which is optimal for fixed r is an open problem.

Yet, we can still give some deviation result. Let α_n and β_n be defined as follows:

$$\alpha_n = n^2\, (p_n)^{4h-3}\, (1-p_n)^3;$$
$$\beta_n = n^{k-2}\, (p_n)^{1-h}\, (1-p_n)^{-1}.$$

With these choices, one has

$$\left|\kappa^{(r)}(X_\gamma^{(n)})\right| \leq (C^4\, r)^r \alpha_n (\beta_n)^r.$$

Unfortunately the convergence speed hypotheses (29) are not always satisfied, so one cannot apply Corollary 3.6 in general. However, one can see that $\kappa^{(2)}(X_\gamma^{(n)})$ and $\kappa^{(3)}(X_\gamma^{(n)})$ are polynomials in n and p_n of degree $2h$ and $3h$ in p_n. Thanks to this observation, if p_n is of order $n^{-\varepsilon}$ with $0 < \varepsilon < 1/6h$, then Conditions (29) are satisfied and Theorem 7.1 still holds in that case.

8. Random character values from central measures on partitions

The combinatorial techniques introduced in Sections 6 and 7 can also be used to study certain models of random integer partitions. Recall that if G is a finite group and if τ is a function $G \to \mathbb{C}$ with $\tau(e_G) = 1$ and $\tau(gh) = \tau(hg)$ (a *trace* on the group), then τ can be expanded uniquely as a linear combination of normalized irreducible characters:

$$\tau = \sum_{\lambda \in \widehat{G}} \mathbb{P}_\tau[\lambda]\, \widehat{\chi}^\lambda.$$

This yields a probability measure $\mathbb{P}_\tau$ on the finite set $\widehat{G}$ of isomorphism classes of irreducible representations of G. This *spectral measure* is *a priori* complex-valued. When $G = \mathfrak{S}(n)$ is the symmetric group of order n, the irreducible representations are labelled by *integer partitions* of size n, that is non-increasing sequences of positive integers

$\lambda = (\lambda_1, \ldots, \lambda_\ell)$ with $\sum_{i=1}^{\ell} \lambda_i = n$. In the following we denote $\mathfrak{P}(n) = \widehat{\mathfrak{S}}(n)$ the set of integer partitions of size n, and $\ell(\lambda) = \ell$ the length of a partition.

Definition 8.1. *A* central measure *on partitions is a family* $(\mathbb{P}_{\tau,n})_{n\in\mathbb{N}}$ *of spectral measures on the sets* $\mathfrak{P}(n)$ *that come from the same trace of the infinite symmetric group* $\mathfrak{S}(\infty)$*. Thus, there exists a trace* $\tau : \mathfrak{S}(\infty) \to \mathbb{C}$ *such that*

$$\tau_{|\mathfrak{S}(n)} = \sum_{\lambda \in \mathfrak{P}(n)} \mathbb{P}_{\tau,n}[\lambda]\, \widehat{\chi}^{\lambda}.$$

Example 8.2. The regular trace $\tau(\sigma) = 1_{\sigma=\mathrm{id}}$ corresponds to the Plancherel measures of the symmetric groups, given by the formula $\mathbb{P}_n[\lambda] = (\dim V^\lambda)^2/n!$, where V^λ is the $\mathfrak{S}(n)$-irreducible module of label λ. They have been extensively studied in connection with Ulam's problem of the longest increasing subsequence and with random matrix theory, see *e.g.* [BDJ99, BOO00, Oko00, IO02].

A central measure $(\mathbb{P}_{\tau,n})_{n\in\mathbb{N}}$ is non-negative if and only if $(\tau(\rho_i\rho_j^{-1}))_{1\le i,j\le n}$ is hermitian non-negative definite for any finite family of permutations $\rho_1, \ldots, \rho_n$. The set of non-negative central measures, *i.e.*, coherent systems of probability measures on partitions has been identified in [Tho64] and later studied in [KV81]. Call extremal a non-negative trace on $\mathfrak{S}_\infty$ that is not a positive linear combination of non-negative traces. Then, extremal central measures are labelled by the infinite-dimensional *Thoma simplex*

$$\Omega = \left\{ \omega = (\alpha, \beta) = \big((\alpha_1 \ge \alpha_2 \ge \cdots \ge 0), (\beta_1 \ge \beta_2 \ge \cdots \ge 0)\big) \;\middle|\; \sum_{i=1}^{\infty} \alpha_i + \beta_i \le 1 \right\}.$$

The trace on the infinite symmetric group corresponding to a parameter ω is given by

$$\tau_\omega(\sigma) = \prod_{c\in C(\sigma)} p_{|c|}(\omega) \quad \text{with } p_1(\omega) = 1,\;\; p_{k\ge 2}(\omega) = \sum_{i=1}^{\infty} (\alpha_i)^k + (-1)^{k-1}(\beta_i)^k,$$

$C(\sigma)$ denoting the set of cycles of σ. In particular, if ρ and ν are two permutations with disjoint non-trivial cycles, then for any extremal trace on $\mathfrak{S}(\infty)$,

$$\tau_\omega(\rho\nu) = \tau_\omega(\rho)\, \tau_\omega(\nu).$$

This property is known as *asymptotic factorization*, and it is involved in most of the asymptotic results on central measures. Kerov and Vershik have shown that if $\omega \in \Omega$ and $\rho \in \mathfrak{S}(\infty)$ are fixed, then the random character value $\widehat{\chi}^\lambda(\rho)$ with λ chosen according to the central measure $\mathbb{P}_{\omega,n}$ converges in probability towards the trace $\tau_\omega(\rho)$. Thus, central measures on partitions and extremal traces of $\mathfrak{S}(\infty)$ are concentrated. More recently, it was shown by Féray and Méliot that this concentration is Gaussian, see [FM12, Mél12]. The aim of this Section is to use the techniques of §6-7 in order to prove the following:

Theorem 8.3. *Fix a parameter* $\omega \in \Omega$*, and a permutation* $\rho \in \mathfrak{S}(\infty)$*. Denote* $X_\rho^{(n)}$ *the random character value* $\widehat{\chi}^\lambda(\rho)$*, where:*

- $\lambda \in \mathfrak{P}(n)$ *is picked randomly according to the central measure* $\mathbb{P}_{\omega,n}$*,*
- *and* $\widehat{\chi}^\lambda$ *denotes the normalized irreducible character of label* λ *of* $\mathfrak{S}(n)$*, that is*

$$\widehat{\chi}^\lambda(\rho) = \frac{\operatorname{tr} \rho^\lambda(\rho)}{\dim V^\lambda} \qquad \textit{with } (V^\lambda, \rho^\lambda) \textit{ irreducible representation of } \mathfrak{S}(n).$$

The rescaled random variable $n^{2/3}\,(X^{(n)}_\rho - \tau_\omega(\rho))$ converges in the mod-Gaussian sense with parameters $t_n = n^{1/3}\,\sigma^2$ and limiting function $\psi(z) = \exp(L\,\frac{z^3}{6})$, where σ^2 and L can be computed explicitly in terms of ω and ρ. In particular, if ρ is a k-cycle, then

$$\sigma^2 = k^2\left(p_{2k-1}(\omega) - p_k(\omega)^2\right);$$
$$L = k^3\left((3k-2)\,p_{3k-2}(\omega) - (6k-3)\,p_{2k-1}(\omega)\,p_k(\omega) + (3k-1)\,p_k(\omega)^3\right).$$

Corollary 8.4. *The random character value $X^{(n)}_k = X^{(n)}_\rho$ with ρ a k-cycle satisfies the principle of moderate deviations*

$$\mathbb{P}\left[|X^{(n)}_k - p_k(\omega)| \geq n^{-1/2}x\right] = \frac{2\,\mathrm{e}^{-\frac{v^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2x^2}}\,\cosh\left(\frac{L\,x^3}{6\,\sigma^3\,n^{1/2}}\right)(1+o(1)),$$

for $1 \ll x \ll n^{1/4}$, with σ^2 and L as in Theorem 8.3.

Actually we shall see during the proof that there is mod-Gaussian convergence for any finite vector of rescaled random character values, so the results of Section 4 holds. This Section is organized as follows. In §8.1, we introduce renormalized conjugacy classes and their cumulants, and we show that these cumulants are polynomials in n. In §8.2, we then prove bounds on these cumulants similar to those of Proposition 7.3, and we compute the limits of the second and third cumulants. This will allow us to use in §8.3 the framework of §3.2 in order to prove the results stated above. We shall also detail some consequences of these results for the shapes of the random partitions $\lambda \sim \mathbb{P}_{n,\omega}$ viewed as Young diagrams, in the spirit of [FM12, Mél12].

8.1. **Renormalized conjugacy classes.** In the whole Section, $\omega \in \Omega$ is fixed. Given a partition $\mu = (\mu_1, \ldots, \mu_\ell)$ of size $|\mu| = \sum_{i=1}^{\ell} \mu_i = k$, we denote

$$\varSigma_{\mu,n} = \sum (a_{1,1}, \ldots, a_{1,\mu_1})(a_{2,1}, \ldots, a_{2,\mu_2}) \cdots (a_{\ell,1}, \ldots, a_{\ell,\mu_\ell}),$$

where the formal sum is taken over arrangements in $\mathfrak{A}(n,k)$ and is considered as an element of the group algebra $\mathbb{C}\mathfrak{S}(n)$. For $A \in \mathfrak{A}(n,k)$, it will be convenient to denote $\rho_\mu(A)$ the corresponding term in $\varSigma_{\mu,n}$.

If $n \geq k$, then $\varSigma_{\mu,n}$ is a multiple of the conjugacy class that consists in elements of cycle type $(\mu_1, \ldots, \mu_\ell, 1^{n-k})$, and it contains $n^{\downarrow k}$ elements; otherwise it is equal to 0. We shall need the following facts about these renormalized character values, for which we refer to [IK99, Śni06b, FM12]:

(1) In the center of the group algebra $\mathbb{C}\mathfrak{S}(n)$, a polynomial in elements $\varSigma_{\mu,n}$ is a linear combination with integer coefficients of $\varSigma_{\nu,n}$'s, and the coefficients of the expansion are independent of n. This can be seen by constructing a projective limit of the algebra of symbols $\varSigma_{\mu,n}$, see [IK99]. An explicit rule to compute products of symbols $\varSigma_{\mu,n}$ is given in [FM12, §3.4].

(2) For any (random) partition $\lambda \in \mathfrak{P}(n)$, we can consider the random variable
$$\varSigma_\mu(\lambda) = \widehat{\chi}^\lambda(\varSigma_{\mu,n}) = n^{\downarrow|\mu|}\,X^{(n)}_{\rho_\mu},$$
where ρ_μ is any permutation of cycle type μ (assuming $k \leq n$). For any polynomial $P(\varSigma_{\mu^{(1)},n}, \ldots, \varSigma_{\mu^{(r)},n})$, it is the same to consider the random variable
$$P(\varSigma_{\mu^{(1)},n}(\lambda), \ldots, \varSigma_{\mu^{(r)},n}(\lambda)),$$

or to first expand the polynomial $P(\Sigma_{\mu^{(1)},n},\dots,\Sigma_{\mu^{(r)},n}) = \sum_\nu c_\nu\, \Sigma_{\nu,n}$ in $\mathbb{C}\mathfrak{S}(n)$, and then evaluate

$$\sum_\nu c_\nu\, \Sigma_{\nu,n}(\lambda).$$

(3) Notice that the expectation under $\mathbb{P}_{n,\omega}$ of a random character value $\widehat{\chi}^\lambda(\rho)$ is $\tau_\omega(\rho)$. For permutations $\rho_1,\rho_2,\dots,\rho_r$ in $\mathfrak{S}(n)$, we define:

$$\kappa(\rho_1,\dots,\rho_r) = \sum_\pi \mu(\pi) \prod_{B\in\pi} \tau_\omega\left(\overrightarrow{\prod_{i\in B}} \rho_i\right),$$

where the sum runs over set-partitions π of $[r]$. Note that the last product is not commutative. By convention, we always multiply the permutations ρ_i by increasing order of their index. This choice is recorded by the arrow above the product sign. The previous remarks show that if $\mu^1,\dots,\mu^r$ are fixed partitions of respective sizes $k_1,\dots,k_r$, then

$$\begin{aligned}
n^{\downarrow k_1}\cdots n^{\downarrow k_r}\ \kappa\left(X^{(n)}_{\rho_{\mu^1}},\dots,X^{(n)}_{\rho_{\mu^r}}\right) &= \kappa(\Sigma_{\mu^1},\dots,\Sigma_{\mu^r})\\
&= \sum_{\pi\in\mathfrak{Q}(r)} \mu(\pi) \prod_{B\in\pi} \tau_\omega\left(\overrightarrow{\prod_{i\in B}} \Sigma_{\mu^i,n}\right)\\
&= \sum_{\substack{A^1\in\mathfrak{A}(n,k_1)\\ \vdots \\ A^r\in\mathfrak{A}(n,k_r)}} \sum_{\pi\in\mathfrak{Q}(r)} \mu(\pi) \prod_{B\in\pi} \tau_\omega\left(\overrightarrow{\prod_{i\in B}} \rho_{\mu^i}(A^i)\right)\\
&= \sum_{\substack{A^1\in\mathfrak{A}(n,k_1)\\ \vdots \\ A^r\in\mathfrak{A}(n,k_r)}} \kappa\left(\rho_{\mu^1}(A^1),\dots,\rho_{\mu^r}(A^r)\right). \qquad (30)
\end{aligned}$$

On the second line of this list of equalities, one has a polynomial (with constant coefficients) in traces $\tau_\omega(\Sigma_{\nu,n})$, because each product $\overrightarrow{\prod}_{i\in B} \Sigma_{\mu^i,n}$ can be expanded as a sum of symbols $\Sigma_{\nu,n}$. However, for every ν, $\tau_\omega(\Sigma_{\nu,n}) = n^{\downarrow|\nu|}\, p_\nu(\omega)$ is a polynomial in n, so:

Lemma 8.5. *Fix some integer partitions $\mu^1,\dots,\mu^r$. Then the rescaled joint cumulant*

$$\kappa(\Sigma_{\mu^1},\dots,\Sigma_{\mu^r}) = n^{\downarrow k_1}\cdots n^{\downarrow k_r}\ \kappa\left(X^{(n)}_{\rho_{\mu^1}},\dots,X^{(n)}_{\rho_{\mu^r}}\right)$$

is a polynomial in n.

This is the analogue of Corollary 7.6. Using now the last line of the list of equalities (30), we are going to bound these joint cumulants.

8.2. **Bounds and limits of the cumulants.**

8.2.1. *The dependency graph of the random character values.* In a noncommutative algebra $\mathscr{A}$ endowed with a trace τ, recall that subalgebras $\mathscr{A}_\lambda$ are said independent if they commute and if

$$\tau(a_1a_2\dots a_k) = \tau(a_1)\tau(a_2)\cdots\tau(a_k)$$

when $a_i\in\mathscr{A}_{\lambda_i}$ with $\lambda_i\neq\lambda_j$ for $i\neq j$. That said, consider a non-commutative algebra $\mathscr{A}$ endowed with a trace $\mathbb{E}:\mathscr{A}\to\mathbb{R}$ that satisfies Equation (22) for a block of variables

B that can be split in two blocks B_1 and B_2 of independent non-commutative random variables. The reader can check that the reasoning of Section 6 can be performed with any family of random variables in $\mathscr{A}$. In other words, one can work with abstract random variables in a non-commutative probability space, that are defined without an underlying notion of randomness; and the estimates on cumulants will still hold. We shall therefore work with the non-commutative probability space $(\mathbb{C}\mathfrak{S}(n), \tau_\omega)$, that indeed satisfies (22). If $k = |\mu|$, we are interested in the non-commutative random variables

$$\Sigma_{\mu,n} = \sum_{A \in \mathfrak{A}(n,k)} \rho_\mu(A).$$

Again, to control the cumulants, we shall exhibit a dependency graph for the families of random variables $\{\rho_\mu(A)\}_{A\in\mathfrak{A}(n,k)}$. By the property of asymptotic factorization, if A and A' are disjoint, then $\rho_\mu(A)$ and $\rho_\mu(A')$ are independent (in the non-commutative sense, w.r.t. the trace τ_ω). Therefore, one can associate to $\{\rho_\mu(A)\}_{A\in\mathfrak{A}(n,k)}$ the dependency graph $\mathbb{B}$ defined by:

- its vertex set is $\mathfrak{A}(n, k)$;
- there is an edge between A and A' if $|A \cap A'| \geq 1$.

The graph $\mathbb{B}$ is obviously regular with degree strictly smaller than $k^2\, n^{\downarrow k-1}$. On the other hand, the expectations of its vertices are all smaller than 1 in absolute value. So one can once again apply Theorem 6.3 and we get:

Proposition 8.6. *Fix a partition μ of size k. For any $r \leq 1$, one has*

$$\left|\kappa^{(r)}(\Sigma_\mu)\right| \leq 2^{r-1} r^{r-2}\, n^{\downarrow k}\, (k^2 n^{\downarrow k-1})^{r-1};$$
$$\left|\kappa^{(r)}(X^{(n)}_{\rho_\mu})\right| \leq r^{r-2} \left(\frac{2k^2}{n}\right)^{r-1}.$$

The multi-variate version of this reads then as:

Proposition 8.7. *Let $\mu^1, \ldots, \mu^r$ be integer partitions of sizes $k_1, \ldots, k_r$. The following bound holds:*

$$\left|\kappa\left(X^{(n)}_{\rho_{\mu^1}}, \ldots, X^{(n)}_{\rho_{\mu^r}}\right)\right| \leq k_1 \cdots k_r\, \left(k_1 + \cdots + k_r\right)^{r-2} \left(\frac{2}{n}\right)^{r-1}.$$

8.2.2. *Limits of the second and third cumulants.* Because of Lemma 8.5 and Proposition 8.7, for any fixed integer partitions,

$$\kappa\left(X^{(n)}_{\rho_{\mu^1}}, \ldots, X^{(n)}_{\rho_{\mu^r}}\right)\, n^{1-r} \simeq \kappa\left(\Sigma_{\mu^1}, \ldots, \Sigma_{\mu^r}\right)\, n^{k_1+\cdots+k_r-(r-1)}$$

converges to a constant. Let us compute this limit when $r = 2$ or 3; we use the same reasoning as in §7.3.2.

As κ is invariant by simultaneous conjugacy of its arguments, the summand in Equation (30) depends only on the graph $G = G_{\mathbb{A}}$ associated to the collection $\mathbb{A} = (A^1, \ldots, A^r)$, and we shall denote it $\kappa(G)$. We fix partitions $\mu^1, \ldots, \mu^r$ of respective sizes $k_1, \ldots, k_r$, and write

$$\kappa(\Sigma_{\mu^1}, \ldots, \Sigma_{\mu^r}) = \sum_G \kappa(G)\, N_G.$$

When $r = 2$, we have to look for graphs G on vertex set $V_{\Bbbk} = [k_1] \sqcup [k_2]$ with 1-contraction connected and at least $k_1 + k_2 - 1$ connected components, because these are the ones that will give a contribution for the coefficient of $n^{k_1+k_2-1}$. For $i \in [\ell(\mu^1)]$ and $j \in [\ell(\mu^2)]$, denote

$$(\mu^1 \bowtie \mu^2)(i,j) = (\mu^1 \setminus \mu^1_i) \sqcup (\mu^2 \setminus \mu^2_j) \sqcup \{\mu^1_i + \mu^2_j - 1\}.$$

This is the cycle type of a permutation $\rho_{\mu^1}(A^1)\,\rho_{\mu^2}(A^2)$, where $G_{\mathbb{A}}$ is the graph with one edge joining an element of A^1 in the cycle of length μ^1_i with an element of A^2 in the cycle of length μ^2_j. These graphs are the only ones involved in our computation, and they yield

$$\kappa(G) = p_{(\mu^1 \bowtie \mu^2)(i,j)}(\omega) - p_{\mu^1 \sqcup \mu^2}(\omega),$$

where for a partition μ we denote $p_\mu(\omega)$ the product $\prod_{i=1}^{\ell(\mu)} p_{\mu_i}(\omega)$. So,

Proposition 8.8. *For any partitions μ and ν, the limit of $n\,\kappa(X^{(n)}_{\rho_\mu}, X^{(n)}_{\rho_\nu})$ is*

$$\sum_{i=1}^{\ell(\mu)} \sum_{j=1}^{\ell(\nu)} \mu_i\,\nu_j\,\left(p_{(\mu \bowtie \nu)(i,j)}(\omega) - p_{\mu \sqcup \nu}(\omega)\right).$$

In particular, for cycles $\mu = (k)$ and $\nu = (l)$,

$$\lim_{n\to\infty} n\,\kappa\left(X^{(n)}_k, X^{(n)}_l\right) = kl\,\left(p_{k+l-1}(\omega) - p_{k,l}(\omega)\right).$$

When $r = 3$, we look for graphs G on vertex set $V_{\Bbbk} = [k_1] \sqcup [k_2] \sqcup [k_3]$ with 1-contraction connected and at least $k_1 + k_2 + k_3 - 2$ connected components. They are of three kinds:

(1) One cycle in $\rho_{\mu^2}(A^2)$ is connected to two cycles in $\rho_{\mu^1}(A^1)$ and $\rho_{\mu^3}(A^3)$, but not by the same point in this cycle of $\rho_{\mu^2}(A^2)$. This gives for $\rho_{\mu^1}(A^1)\,\rho_{\mu^2}(A^2)\,\rho_{\mu^3}(A^3)$ a permutation of cycle type

$$(\mu^1 \bowtie \mu^2 \bowtie \mu^3)(i,j,k) = (\mu^1 \setminus \mu^1_i) \sqcup (\mu^2 \setminus \mu^2_j) \sqcup (\mu^3 \setminus \mu^3_k) \sqcup \{\mu^1_i + \mu^2_j + \mu^3_k - 2\};$$

and the corresponding cumulant is

$$\kappa(G) = p_{\mu^1 \sqcup \mu^2 \sqcup \mu^3}(\omega) + p_{(\mu^1 \bowtie \mu^2 \bowtie \mu^3)(i,j,k)}(\omega) - p_{((\mu^1 \bowtie \mu^2)(i,j)) \sqcup \mu^3}(\omega) - p_{((\mu^2 \bowtie \mu^3)(j,k)) \sqcup \mu^1}(\omega).$$

In this description, one can permute cyclically the indices $1, 2, 3$, and this gives 3 different graphs.

(2) One cycle in $\rho_{\mu^2}(A^2)$ is connected to two cycles in $\rho_{\mu^1}(A^1)$ and $\rho_{\mu^3}(A^3)$, and by the same point in this cycle of $\rho_{\mu^2}(A^2)$. In other words, there is an identity $a^1_s = a^2_t = a^3_u$. This gives again for $\rho_{\mu^1}(A^1)\,\rho_{\mu^2}(A^2)\,\rho_{\mu^3}(A^3)$ a permutation of cycle type $(\mu^1 \bowtie \mu^2 \bowtie \mu^3)(i,j,k)$, but the corresponding cumulant takes now the form

$$\kappa(G) = 2\,p_{\mu^1 \sqcup \mu^2 \sqcup \mu^3}(\omega) + p_{(\mu^1 \bowtie \mu^2 \bowtie \mu^3)(i,j,k)}(\omega) - p_{((\mu^1 \bowtie \mu^2)(i,j)) \sqcup \mu^3}(\omega) - p_{((\mu^2 \bowtie \mu^3)(j,k)) \sqcup \mu^1}(\omega) - p_{((\mu^1 \bowtie \mu^3)(i,k)) \sqcup \mu^2}(\omega).$$

Here, there is no need to permute cyclically the indices in the enumeration for N_G.

(3) Two distinct cycles in $\rho_{\mu^2}(A^2)$ are connected to a cycle of $\rho_{\mu^1}(A^1)$ and to a cycle of $\rho_{\mu^3}(A^3)$, which gives a permutation of cycle type

$$(\mu^1 \bowtie \mu^2 \bowtie \mu^3)(i,j;k,l) = (\mu^1 \setminus \mu^1_i) \sqcup (\mu^2 \setminus \{\mu^2_j, \mu^2_k\}) \sqcup (\mu^3 \setminus \mu^3_l) \sqcup \{\mu^1_i + \mu^2_j - 1, \mu^2_k + \mu^3_l - 1\}.$$

The cumulant corresponding to this last case is

$$\kappa(G) = p_{\mu^1\sqcup\mu^2\sqcup\mu^3}(\omega) + p_{(\mu^1\bowtie\mu^2\bowtie\mu^3)(i,j;k,l)}(\omega) - p_{((\mu^1\bowtie\mu^2)(i,j))\sqcup\mu^3}(\omega) - p_{((\mu^2\bowtie\mu^3)(k,l))\sqcup\mu^1}(\omega),$$

and again one can permute cyclically the indices $1, 2, 3$ to get 3 different graphs.

Consequently:

Proposition 8.9. *For any partitions μ, ν and δ, the limit of $n^2\,\kappa(X^{(n)}_{\rho_\mu}, X^{(n)}_{\rho_\nu}, X^{(n)}_{\rho_\delta})$ is*

$$\sum_{\mathbb{Z}/3\mathbb{Z}}\left(\sum_{i=1}^{\ell(\mu)}\sum_{j=1}^{\ell(\nu)}\sum_{k=1}^{\ell(\delta)} \mu_i\,\nu_j\,(\nu_j-1)\,\delta_k\left(\begin{smallmatrix} p_{\mu\sqcup\nu\sqcup\delta}(\omega)+p_{(\mu\bowtie\nu\bowtie\delta)(i,j,k)}(\omega)\\ -p_{((\mu\bowtie\nu)(i,j))\sqcup\delta}(\omega)-p_{((\nu\bowtie\delta)(j,k))\sqcup\mu}(\omega)\end{smallmatrix}\right)\right.$$
$$\left.+\sum_{i=1}^{\ell(\mu)}\sum_{(j\neq k)=1}^{\ell(\nu)}\sum_{l=1}^{\ell(\delta)} \mu_i\,\nu_j\,\nu_k\,\delta_l\left(\begin{smallmatrix} p_{\mu\sqcup\nu\sqcup\delta}(\omega)+p_{(\mu\bowtie\nu\bowtie\delta)(i,j;k,l)}(\omega)\\ -p_{((\mu\bowtie\nu)(i,j))\sqcup\delta}(\omega)-p_{((\nu\bowtie\delta)(k,l))\sqcup\mu}(\omega)\end{smallmatrix}\right)\right)$$
$$+\sum_{i=1}^{\ell(\mu)}\sum_{j=1}^{\ell(\nu)}\sum_{k=1}^{\ell(\delta)} \mu_i\,\nu_j\,\delta_k\left(\begin{smallmatrix} 2\,p_{\mu\sqcup\nu\sqcup\delta}(\omega)+p_{(\mu\bowtie\nu\bowtie\delta)(i,j,k)}(\omega)-p_{((\mu\bowtie\nu)(i,j))\sqcup\delta}(\omega)\\ -p_{((\nu\bowtie\delta)(j,k))\sqcup\mu}(\omega)-p_{((\mu\bowtie\delta)(i,k))\sqcup\nu}(\omega)\end{smallmatrix}\right),$$

where $\sum_{\mathbb{Z}/3\mathbb{Z}}$ means that one permutes cyclically the partitions μ, ν, δ. In particular, for cycles $\mu = (k)$, $\nu = (l)$ and $\delta = (m)$, $\lim_{n\to\infty} n^2\,\kappa(X^{(n)}_k, X^{(n)}_l, X^{(n)}_m)$ is equal to

$$klm\big((k+l+m-1)\,p_{k,l,m}(\omega) + (k+l+m-2)\,p_{k+l+m-2}(\omega) - (k+l-1)\,p_{k+l-1,m}(\omega) - (l+m-1)\,p_{l+m-1,k}(\omega) - (l+m-1)\,p_{k+m-1,l}(\omega)\big).$$

One recovers for $k = l = m$ the values of σ^2 and L announced in Theorem 8.3.

8.3. **Asymptotics of the random character values and partitions.** Fix integer partitions $\mu^1, \ldots, \mu^r$, and consider the random vector

$$\mathbf{V} = n^{2/3}\,(X^{(n)}_{\rho_{\mu^1}} - p_{\mu^1}(\omega), \ldots, X^{(n)}_{\rho_{\mu^r}} - p_{\mu^r}(\omega)).$$

The previous results show that the asymptotics of the multivariate generating series are

$$\mathbb{E}[\mathrm{e}^{\langle \mathbf{V}\,|\,\mathbf{z}\rangle}] = \exp\left(\frac{n^{1/3}}{2}\sum_{i,j=1}^{r}\kappa^{(i,j)}(\mu^1,\ldots,\mu^r)\,\mathbf{z}_i\mathbf{z}_j + \frac{1}{6}\sum_{i,j,k}\kappa^{(i,j,k)}(\mu^1,\ldots,\mu^r)\,\mathbf{z}_i\mathbf{z}_j\mathbf{z}_k\right),$$

where $\kappa^{(i,j)} = \lim_{n\to\infty} n\,\kappa(X^{(n)}_{\rho_{\mu^1}}, X^{(n)}_{\rho_{\mu^2}})$ and $\kappa^{(i,j,k)} = \lim_{n\to\infty} n^2\,\kappa(X^{(n)}_{\rho_{\mu^1}}, X^{(n)}_{\rho_{\mu^2}}, X^{(n)}_{\rho_{\mu^3}})$ are the limiting quantities given in §8.2.2. So, one has mod-Gaussian convergence for every vector of random character values, and this ends the proof of Theorem 8.3 and Corollary 8.4. Note that the speed of convergence of the cumulants is each time a $O((\alpha_n)^{-1})$ because of the polynomial behavior established in Lemma 8.5; therefore, one can indeed apply Corollary 3.6 with $\alpha_n = n$ and $\beta_n = n^{-1}$.

These estimates on the distribution of the random character values are sharp but in one case: when $\omega = ((0,0,\ldots),(0,0,\ldots))$. This parameter of the Thoma simplex corresponds to the Plancherel measures of the symmetric groups, and in this case, since $p_2(\omega) = p_3(\omega) = \cdots = 0$, the parameters of the mod-Gaussian convergence are all equal to 0. Indeed, the random character values under Plancherel measures do not have fluctuations

of order $n^{-1/2}$. For instance, Kerov's central limit theorem (*cf.* [IO02]) ensures that the random character values

$$\frac{n^{k/2}\,\widehat{\chi}^{\lambda}(c_k)}{\sqrt{k}}$$

on cycles c_k of lengths $k \geq 2$ converges in law towards independent Gaussian variables; so the fluctuations are of order $n^{-k/2}$ instead of $n^{-1/2}$. One still expects a mod-Gaussian convergence for adequate renormalizations of the random character values; however, the combinatorics underlying the asymptotics of Plancherel measures are much more complex than those of general central measures, see [Śni06a].

From the estimates on the laws of the random character values, one can prove many estimates for the parts $\lambda_1, \lambda_2, \ldots$ of the random partitions taken under central measures. The arguments of algebraic combinatorics involved in these deductions are detailed in [FM12, Mél12], so here we shall only state results. Given a partition $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ of size n, the Frobenius coordinates of λ are the two sequences

$$\left(\lambda_1 - \frac{1}{2}, \lambda_2 - \frac{3}{2}, \ldots, \lambda_d - d + \frac{1}{2}\right), \left(\lambda_1' - \frac{1}{2}, \lambda_2' - \frac{3}{2}, \ldots, \lambda_d' - d + \frac{1}{2}\right)$$

where λ_1', λ_2', *etc.* are the sizes of the columns of the Young diagram of λ, and d is the size of the diagonal of the Young diagram. Denote $(a_1, \ldots, a_d), (b_1, \ldots, b_d)$ these coordinates, and

$$X_\lambda = \sum_{i=1}^{d} \frac{a_i}{n}\,\delta_{\left(\frac{a_i}{n}\right)} + \sum_{i=1}^{d} \frac{b_i}{n}\,\delta_{\left(-\frac{b_i}{n}\right)}.$$

This is a (random) discrete probability measure on $[-1, 1]$ whose moments

$$p_k(\lambda) = n^k \int_{-1}^{1} x^{k-1}\,X_\lambda(dx)$$

are also the moments of the Frobenius coordinates, so X_λ encodes the geometry of the Young diagram λ. We shall also need

$$X_\omega = \sum_{i=1}^{d} \alpha_i\,\delta_{(\alpha_i)} + \sum_{i=1}^{d} \beta_i\,\delta_{(-\beta_i)} + \gamma\,\delta_{(0)},$$

which will appear in a moment as the limit of the random measures X_λ. Here, $\gamma = 1 - \sum_{i=1}^{\infty} \alpha_i - \sum_{i=1}^{\infty} \beta_i$. Notice that $\mathbb{E}_{n,\omega}[X_k^{(n)}] = \tau_\omega(c_k) = X_\omega(x^{k-1})$.

It is shown in [IO02] that for any partition λ of size n and for any k,

$$p_k(\lambda) = \varSigma_k(\lambda) + \text{remainder},$$

where the remainder is a linear combination of symbols $\varSigma_\mu$ with $|\mu| < k$. It follows that the cumulants of the p_k's satisfy the same estimates as the cumulants of the $\varSigma_k$'s. Therefore, the rescaled random variable $\nabla(x^{k-1}) = n^{2/3}\,(X_\lambda(x^{k-1}) - X_\omega(x^{k-1}))$ converges in the mod-Gaussian sense with parameters $n^{1/3}\,\sigma^2$ and limiting function $\psi(z) = \exp(L\,\frac{z^3}{6})$, where σ^2 and L are given by the same formula as in the case of the random character value $X_k^{(n)}$, that is to say

$$\sigma^2 = k^2\left(p_{2k-1}(\omega) - p_k(\omega)^2\right);$$
$$L = k^3\left((3k-2)\,p_{3k-2}(\omega) - (6k-3)\,p_{2k-1}(\omega)\,p_k(\omega) + (3k-1)\,p_k(\omega)^3\right).$$

Actually, one has mod-Gaussian convergence for any finite vector of random variables $\nabla(x^{k-1})$, $k \geq 1$.

Now, there is a convenient way to rewrite the limiting values of the second and third cumulants. Denote $T : \mathbb{C}[x] \to \mathbb{C}[x]$ the operator

$$P(x) \mapsto \frac{\partial}{\partial x}(xP(x)).$$

Then, for any polynomial P, if $\nabla(P(x)) = n^{2/3}\left(X_\lambda(P(x)) - X_\omega(P(x))\right)$, then $\nabla(P(x))$ converges in the mod-Gaussian sense with parameters

$$t_n = n^{1/3}\,\left(X_\omega((TP)^2) - X_\omega(TP)\,X_\omega(TP)\right) \tag{31}$$

and limiting function $\exp(L\,\frac{z^3}{6})$, with L equal to

$$X_\omega(T((TP)^3)) - (X_\omega(TP))^3 - 3\,X_\omega(TP)\,X_\omega(T((TP)^2)) + 3\,X_\omega(T^2P)\,(X_\omega(TP))^2. \tag{32}$$

The same holds for vectors of fluctuations $(\nabla(P_1(x)), \ldots, \nabla(P_r(x)))$, with limiting joint cumulants obtained from Equations (31) and (32) by "polarization": for instance, the limiting joint covariances are $X_\omega(TP_{i_1}\,TP_{i_2}) - X_\omega(TP_{i_1})\,X_\omega(TP_{i_2})$.

All this leads to the following result of mod-Gaussian convergence, which has a infinite-dimensional flavour. It is convenient to introduce

$$\begin{aligned} \Delta^{(k)} : \mathbb{C}[x] &\to \mathbb{C}[x]^{\otimes k} \\ P &\mapsto P^{\otimes k}; \end{aligned}$$

and for an operator $O : \mathbb{C}[x] \to \mathbb{C}[x]$ or a linear form $F : \mathbb{C}[x] \to \mathbb{C}$,

$$\begin{aligned} O^{(\star)} : \bigoplus_{k=0}^{\infty} \mathbb{C}[x]^{\otimes k} &\to \mathbb{C}[x] \\ P_1 \otimes P_2 \otimes \cdots \otimes P_k &\mapsto O(P_1) \times O(P_2) \times \cdots \times O(P_k); \\ F^{(\star)} : \bigoplus_{k=0}^{\infty} \mathbb{C}[x]^{\otimes k} &\to \mathbb{C} \\ P_1 \otimes P_2 \otimes \cdots \otimes P_k &\mapsto F(P_1)\,F(P_2)\cdots F(P_k). \end{aligned}$$

Theorem 8.10. *The fluctuations of random measures $n^{2/3}(X_\lambda - X_\omega)$ converge in the mod-Gaussian sense with covariance operator*

$$\Xi_{n,\omega} = n^{1/3}(X_\omega)^{(\star)} \circ \left(T^{(\star)} \circ \Delta^{(2)} - \Delta^{(2)} \circ T^{(\star)}\right),$$

and limiting functional

$$\Psi = \exp\left(\frac{(X_\omega T)^{(\star)}}{6} \circ \left(T^{(\star)} - 3\,\mathrm{id} \otimes T^{(\star)} + 3\,\mathrm{id}^{\otimes 2} \otimes T^{(\star)} - \mathrm{id}^{\otimes 3}\right) \circ \Delta^{(3)}\right).$$

By that we mean that for any polynomial $P = z_1\,P_1(x) + z_2\,P_2(x) + \cdots + z_r\,P_r(x)$ in $\mathbb{C}[x][z_1, \ldots, z_r]$, one has the asymptotics

$$\mathbb{E}\left[\mathrm{e}^{n^{2/3}(X_\lambda(P) - X_\omega(P))}\right] = \exp\left(\frac{\Xi_n(P)}{2}\right)\Psi(P)\,(1 + o(1)).$$

It is difficult to obtain from there moderate deviations for the parts $\lambda_1, \lambda_2, \ldots$ of the partition; but the theorem gives at least a central limit theorem when one has strict inequalities $\alpha_1 > \alpha_2 > \cdots > \alpha_i > \cdots$ and $\beta_1 > \beta_2 > \cdots > \beta_i > \cdots$ (see [Mél12], and

also [Buf12]). Indeed, replacing the polynomial P by a smooth test function ψ_i equal to 1 around α_i and to 0 outside a neighborhood of this point, one has

$$X_\lambda(\psi_i) - X_\omega(\psi_i) = \frac{a_i}{n} - \alpha_i \tag{33}$$

with probability going to 1, and Theorem 8.10 ensures that the quantities in the left-hand side renomalized by $\sqrt{n}$ converge jointly towards a Gaussian vector with covariance

$$\kappa(i,j) = X_\omega(T\psi_i\, T\psi_j) - X_\omega(T\psi_i)\, X_\omega(T\psi_j) = \delta_{ij}\, \alpha_i - \alpha_i\alpha_j. \tag{34}$$

So, the fluctuations

$$\sqrt{n}\left(\frac{\lambda_i}{n} - \alpha_i\right)$$

of the rows of the random partitions taken under central measures $\mathbb{P}_{n,\omega}$ converge jointly towards a gaussian vector with covariances given by Equation (34), and one can include in this result the fluctuations

$$\sqrt{n}\left(\frac{\lambda_j'}{n} - \beta_j\right)$$

of the columns of the random partitions, with a similar formula for their covariances.

The reason why it becomes difficult to get by the same technique the moderate deviations of the rows and columns is that in Equation (33), one throws away an event of probability going to zero (because of the law of large numbers satisfied by the rows and the columns, see *e.g.* [KV81]). However, one cannot *a priori* neglect this event in comparison to rare events such as $\{a_i - n\alpha_i \geq n^{2/3}\, x\}$; indeed, these rare events are themselves of probability going exponentially fast to zero. Also, there is the problem of approximation of smooth test functions by polynomials, which one has to control precisely when doing these computations. One still conjectures these moderate deviations to hold, and $n^{2/3}\left(\frac{a_i}{n} - \alpha_i\right)$ to converge in the mod-Gaussian sense with parameters $n^{1/3}(\alpha_i - \alpha_i^2)$ and limiting function

$$\psi(z) = \exp\left(\frac{\alpha_i - 3\alpha_i^2 + 2\alpha_i^3}{6}\, z^3\right)$$

— this is what is given by Theorem 8.10 if we ignore the previous caveats, and still suppose the α_i and β_j all distinct. As explained in [Mél12], this would give moderate deviations for the lengths of the longest increasing subsequences in a random permutation obtained by generalized riffle shuffle.

9. Appendices

9.1. **Properties of the kernels $\mu^{(k)}_{T,\mathbb{R}^d}$.** We prove here Lemma 4.3. On the hypercube $[-\varepsilon,\varepsilon]^d$, $|F(\mathbf{x}-\mathbf{y}) - F(\mathbf{x})|$ is smaller than $m\, d^{1/2}\, \varepsilon$, whereas on the complementary, the mass of the integral kernel is smaller than

$$\frac{2d}{I(k)} \int_\varepsilon^\infty \frac{1}{T^{2k+1}\, x^{2k+2}}\, dx = \frac{2d\,(2k+1)}{I(k)\,(T\varepsilon)^{2k+1}}.$$

It can be shown that $I(k)$ is always a rational multiple of π, decreasing with k — for instance, $I(0)=\pi$ and $I(1)=\frac{5\pi}{12}$. On $[-\frac{\pi}{2\sqrt{2k+1}},\frac{\pi}{2\sqrt{2k+1}}]$,

$$\frac{1-\cos(x)}{x^2}\geq \frac{1}{2}-\frac{\pi^2}{96(2k+1)}\geq \frac{1}{2}\left(1-\frac{\pi^2}{24(2k+1)}\right)$$
$$\frac{\sin x}{x}\geq 1-\frac{\pi^2}{24(2k+1)}.$$

So,

$$\begin{aligned}I(k)&\geq \int_{-\frac{\pi}{2\sqrt{2k+1}}}^{\frac{\pi}{2\sqrt{2k+1}}}\left(\frac{1-\cos x}{x^2}\right)\left(\frac{\sin x}{x}\right)^{2k}dx\\ &\geq \frac{\pi}{2\sqrt{2k+1}}\left(1-\frac{\pi^2}{24(2k+1)}\right)^{2k+1}\geq \frac{\pi\left(1-\frac{\pi^2}{24}\right)}{2\sqrt{2k+1}},\end{aligned}$$

and therefore, $\int_{([-\varepsilon,\varepsilon]^d)^c}\Delta^{(k)}_{T,\mathbb{R}^d}(\mathbf{x})\,d\mathbf{x}\leq \frac{5\,d\,(2k+1)^{3/2}}{2(T\varepsilon)^{2k+1}}$. Then,

$$\begin{aligned}|\mu^{(k)}_T(F)-\mu(F)|&\leq \|\Delta^{(k)}_{T,\mathbb{R}^d}*F-F\|_\infty\\ &\leq \sup_{\mathbf{x}\in\mathbb{R}^d}\left(\int_{[-\varepsilon,\varepsilon]^d}+\int_{([-\varepsilon,\varepsilon]^d)^c}\right)\left(|F(\mathbf{x}-\mathbf{y})-F(\mathbf{x})|\,\Delta_{T,\mathbb{R}^d}(\mathbf{y})\,d\mathbf{y}\right)\\ &\leq \inf_{\varepsilon>0}\left(m\,d^{1/2}\,\varepsilon+\frac{5\,C\,d\,(2k+1)^{3/2}}{(T\,\varepsilon)^{2k+1}}\right)\leq \frac{27}{2}\,d^{\frac{2k+3}{4k+4}}\left(\frac{C\,m^{2k+1}}{T^{2k+1}}\right)^{\frac{1}{2k+2}}\end{aligned}$$

by choosing on the last line the minimizer $\varepsilon=\left((5\,C\,d^{1/2}\,(2k+1)^{5/2})/(mT^{2k+1})\right)^{\frac{1}{2k+2}}$ of the previous bound, and simplifying a bit the constants.

The second part of Lemma 4.3 relies on the following arguments. Denote $\mathscr{M}_T$ the operator on $\mathscr{C}^0(\mathbb{R})$ defined by

$$\mathscr{M}_T f(x)=\frac{1}{2T}\int_{x-T}^{x+T}f(y)\,dy.$$

The image of $\mathscr{C}^k(\mathbb{R})$ by $\mathscr{M}_T$ is included into $\mathscr{C}^{k+1}(\mathbb{R})$, and on the other hand, if f is supported by $[a,b]$, then $\mathscr{M}_T f$ is supported by $[a-T,b+T]$. At the level of Fourier transforms, one has:

$$\begin{aligned}(\mathscr{M}_T\widehat{f})(x)&=\frac{1}{2T}\int_{x-T}^{x+T}\widehat{f}(y)\,dy=\frac{1}{2T}\int_{\mathbb{R}}f(z)\left(\int_{x-T}^{x+T}\mathrm{e}^{\mathrm{i}zy}\,dy\right)dz\\ &=\int_{\mathbb{R}}f(z)\,\frac{\sin Tz}{Tz}\,\mathrm{e}^{\mathrm{i}zx}\,dz=\widehat{(fg_T)}(x),\end{aligned}$$

with $g_T(x)=\frac{\sin Tx}{Tx}$. As a consequence, if f is a non-negative function with Fourier transform supported by $[-T,T]$, then $(g_T)^{2k}f$ is a non-negative function with Fourier transform of class $\mathscr{C}^{2k}$ and supported by $[-(2k+1)T,(2k+1)T]$. However,

$$\Delta^{(0)}_T(x)=\frac{1}{\pi}\,\frac{1-\cos Tx}{Tx^2}\quad\Rightarrow\quad\widehat{\Delta^{(0)}_T}(x)=\left(1-\frac{|x|}{T}\right)_+,$$

which is supported by $[-T,T]$ and with values in $[0,1]$. So, the Fourier transform of $\Delta^{(k)}_T$ takes its values in $[0,1]$, is supported by $[-(2k+1)T,(2k+1)T]$ and is of class $\mathscr{C}^{2k}$. The result for $\Delta^{(k)}_{T,\mathbb{R}^d}$ follows immediately.

9.2. **Gaussian regularity of convex bodies.** Let B be a convex body in $\mathbb{R}^d$ containing $\mathbf{0}$ in its interior, homeomorphic to a closed ball and with a boundary ∂B which we assume to be Lipschitz in the following sense: there exists a constant $L > 0$ such that the boundary ∂B is parametrized by an homeomorphism $\chi : \mathbb{S}^{d-1} \to \partial B$ with

$$\chi(\theta) \in \mathbb{R}_+\theta \qquad ; \qquad \frac{\|\chi(\theta) - \chi(\theta')\|}{\|\chi(\theta)\|} \leq L\,\|\theta - \theta'\|.$$

Notice that the constant L does not change if one renormalizes B and looks at λB instead of B, $\lambda > 0$. In other words, L only depends on the form of the convex body viewed from the origin. We refer to [Sch93, Hör94] for references on convex bodies. We claim that under these assumptions, B is Gaussian regular with constant $4(1+2L)^2$. As will be clear from the proof, the origin $\mathbf{0}$ only plays a role of reference point, so the same holds for any convex body with boundary L-Lipschitz viewed from a given point in the interior of B.

First, for any direction $\theta \in \mathbb{S}^{d-1}$, we claim that the intersection of $\mathbb{R}_+\theta$ with $B^\varepsilon \setminus B$ is an interval. Indeed, if two points $\mathbf{a}, \mathbf{b}$ on $\mathbb{R}_+\theta$ are in $B^\varepsilon \setminus B$, then there exists corresponding points $\mathbf{c}, \mathbf{d}$ in B with $\|\mathbf{a} - \mathbf{c}\| \leq \varepsilon$ and $\|\mathbf{b} - \mathbf{d}\| \leq \varepsilon$, and then by convexity of B all the points between $\mathbf{a}$ and $\mathbf{b}$ are in B^ε. So, there exists real numbers $0 < \alpha < \beta$ such that

$$(B^\varepsilon \setminus B) \cap \mathbb{R}_+\theta = (\alpha\theta, \beta\theta],$$

and in fact one has necessarily $\alpha\theta \in \partial B$, and $\alpha\theta = \chi(\theta)$. Then, $\beta\theta$ is ε-close to some $\chi(\theta')$, so

$$\beta - \alpha = \|\beta\theta - \chi(\theta)\| \leq \|\beta\theta - \chi(\theta')\| + \|\chi(\theta') - \chi(\theta)\| \leq \varepsilon + L\,\|\chi(\theta)\|\,\|\theta - \theta'\|.$$

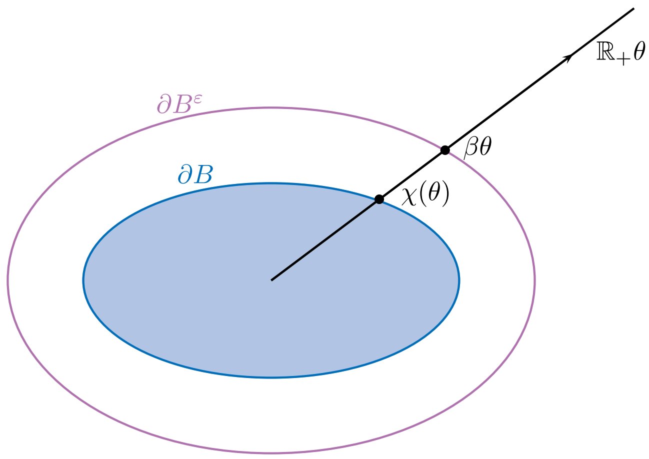


FIGURE 14. The intersection of an half-line $\mathbb{R}_+\theta$ with $B^\varepsilon \setminus B$, B convex body is a segment whose length is bounded by a multiple of ε if $\mathbf{0} \in B^{\mathrm{o}}$ and ∂B is Lipschitz.

However, for two points $\mathbf{a}$ and $\mathbf{b}$ on two lines $\mathbb{R}_+\theta$ and $\mathbb{R}_+\theta'$, one has always $\|\mathbf{a} - \mathbf{b}\| \geq \frac{\|\mathbf{a}\|}{2}\|\theta - \theta'\|$, the worst case being when $\theta = -\theta'$ and $\mathbf{b}$ is near zero. Consequently,

$$\beta - \alpha \leq \varepsilon + 2L\frac{\|\chi(\theta)\|}{\|\beta\theta\|}\,\|\chi(\theta') - \beta\theta\| \leq (2L+1)\,\varepsilon.$$

Then, one writes:

$$\begin{aligned}\frac{1}{(2\pi)^{d/2}}\int_{B^\varepsilon\setminus B}\mathrm{e}^{-\frac{\|\mathbf{x}\|^2}{2}}\,d\mathbf{x} &= \frac{1}{(2\pi)^{d/2}}\int_{\mathbb{S}^{d-1}}\int_{\alpha(\theta)}^{\beta(\theta)}\left(\mathrm{e}^{-\frac{r^2}{2}}\,r^{d-1}\,dr\right)\,d\mu_{\mathbb{S}^{d-1}}(\theta)\\ &\leq \frac{\mathrm{vol}(\mathbb{S}^{d-1})}{(2\pi)^{d/2}}\,\max_{\alpha>0}\left(\int_{\alpha}^{\alpha+(2L+1)\varepsilon}\mathrm{e}^{-\frac{r^2}{2}}\,r^{d-1}\,dr\right)\\ &\leq (2L+1)\,\varepsilon\,\frac{2^{1-d/2}}{\Gamma(d/2)}\,\max_{r\geq 0}\left(\mathrm{e}^{-\frac{r^2}{2}}\,r^{d-1}\right)\leq 2\,(2L+1)\,\varepsilon\end{aligned}$$

by using Stirling estimates.

Similarly, we claim that for any direction $\theta\in\mathbb{S}^{d-1}$, the intersection of $\mathbb{R}_+\theta$ with $B\setminus B^{-\varepsilon}$ is an interval. Fix $\varepsilon,\eta>0$, a direction θ and consider the set

$$I_{\varepsilon,\eta,\theta}=\{\mathbf{x}\in B\cap\mathbb{R}\theta\mid B_{(\mathbf{x},\varepsilon+\eta)}\subset B\}.$$

This is a convex compact subset of $\mathbb{R}\theta$, as shown by Figure 15. However, $B^{-\varepsilon}=((B^{\mathrm{c}})^{\varepsilon})^{\mathrm{c}}=\bigcup_{\eta>0}\{\mathbf{x}\in B\mid B_{(\mathbf{x},\varepsilon+\eta)}\subset B\}$, so $\mathbb{R}\theta\cap B^{-\varepsilon}$ is the increasing union of the segments $I_{\varepsilon,\eta,\theta}$. Therefore, there exists real numbers $0\leq\gamma<\alpha$ such that

$$(B\setminus B^{-\varepsilon})\cap\mathbb{R}_+\theta=[\gamma\theta,\alpha\theta],$$

possibly with $\gamma=0$. As before, $\alpha\theta=\chi(\theta)$. Denote θ' a direction such that $\gamma\theta$ is ε-close to a point of $\mathbb{R}_+\theta'$ which is outside B; one can assume without loss of generality that this point is on the boundary of B, that is to say that $\|\gamma\theta-\chi(\theta')\|\leq\varepsilon$. Again, we want to bound $\alpha-\gamma$ by a constant times ε.

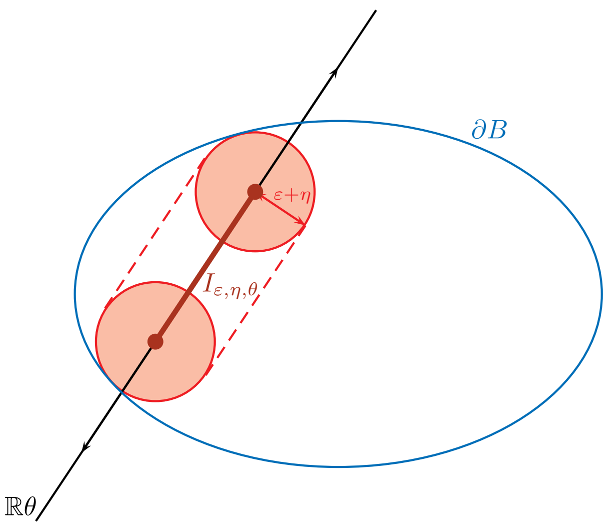


FIGURE 15. The intersection of a line $\mathbb{R}\theta$ with $B^{-\varepsilon}=((B^{\mathrm{c}})^{\varepsilon})^{\mathrm{c}}$ is the union of all sets $I_{\varepsilon,\eta,\theta}$, $\eta>0$, and these sets are nested segments.

- If $\gamma\geq 4L\,\varepsilon$, then one has

$$\alpha-\gamma=\|\chi(\theta)-\gamma\theta\|\leq\|\chi(\theta')-\gamma\theta\|+\|\chi(\theta')-\chi(\theta)\|\leq\varepsilon+L\,\|\chi(\theta)\|\,\|\theta-\theta'\|.$$

The second term in the right-hand side is smaller than $L\,\alpha\,\frac{2}{\gamma}\|\gamma\theta-\chi(\theta')\|\leq\frac{2L\,\varepsilon\,\alpha}{\gamma}$, so

$$\alpha-\gamma\leq\varepsilon+\frac{2L\varepsilon\,\alpha}{\gamma}\qquad;\qquad R\leq 1+\frac{(2L+1)\varepsilon}{\gamma-2L\varepsilon}.$$

with $R = \frac{\alpha}{\gamma}$. Thus, $\alpha \leq \gamma + (2L+1)\varepsilon \, \frac{\gamma}{\gamma - 2L\varepsilon}$, and by the hypothesis $\gamma \geq 4L\,\varepsilon$, this leads to $\alpha - \gamma \leq 2(1+2L)\,\varepsilon$.

- Conversely, if $\gamma \leq 4L\,\varepsilon$, then $\|\chi(\theta')\| \leq (4L+1)\,\varepsilon$, so

$$\begin{aligned}\alpha - \gamma = \|\chi(\theta) - \gamma\theta\| &\leq \|\chi(\theta') - \gamma\theta\| + \|\chi(\theta') - \chi(\theta)\| \leq \varepsilon + L\,\|\chi(\theta')\|\,\|\theta - \theta'\| \\ &\leq (1 + 2L(4L+1))\,\varepsilon \leq (8L^2 + 4L + 2)\,\varepsilon.\end{aligned}$$

  So in any case, $\alpha - \gamma \leq 2(1+2L)^2\,\varepsilon$.

Using then spherical coordinates to compute the integrals, we obtain the bound

$$\frac{1}{(2\pi)^{d/2}} \int_{B \setminus B^{-\varepsilon}} \mathrm{e}^{-\frac{\|\mathbf{x}\|^2}{2}}\,d\mathbf{x} \leq 4(1+2L)^2\,\varepsilon.$$

Hence, we have indeed proved that convex bodies with boundary Lipschitz with constant L viewed from a certain point of their interior are Gaussian-regular with constant $4(1+2L)^2$. Now, it should be noticed that any convex body B with non-empty interior has a Lipschitz boundary for a certain constant L and with respect to a certain reference point $\mathbf{z}$ in B^{o}. However, this constant L may be extremely big if one does not choose $\mathbf{z}$ correctly (*e.g.* near the boundary of B). On the other hand, it follows from [BR10] that one can in fact choose a constant G of Gaussian regularity that works for every convex body of $\mathbb{R}^d$, but the proof of this relies then on more complex arguments than before, and is much longer.

References

[ABT03] R. Arratia, A. D. Barbour, and S. Tavaré. *Logarithmic Combinatorial Structures: a Probabilistic Approach.* EMS Monographs in Mathematics. European Mathematical Society, 2003.

[BDJ99] J. Baik, P. Deift, and K. Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.*, 12:1119–1178, 1999.

[Ben13] C. G. Beneš. A local central limit theorem and loss of rotational symmetry of planar simple random walk. `arXiv:1302.2971v1 [math.PR]`, 2013.

[Bin86] N. H. Bingham. Variants on the law of the iterated logarithm. *Bull. London Math. Soc.*, 18(5):433–467, 1986.

[BKN13] A. D. Barbour, E. Kowalski, and A. Nikeghbali. Mod-discrete expansions, 2013.

[Bol98] B. Bollobás. *Modern graph theory*, volume 184 of *Graduate texts in mathematics*. Springer-Verlag, 1998.

[Bol01] B. Bollobás. *Random graphs*, volume 73 of *Cambridge studies in advanced mathematics*. Cambridge University Press, 2001. Second edition.

[Bón10] M. Bóna. On three different notions of monotone subsequences. In *Permutation Patterns*, volume 376 of *London Math. Soc. Lecture Note Series*, pages 89–113. Cambridge University Press, 2010.

[BOO00] A. Borodin, A. Okounkov, and G. Olshanski. Asymptotics of Plancherel measures for symmetric groups. *J. Amer. Math. Soc.*, 13:491–515, 2000.

[Bor60] C. W. Borchardt. Über eine der Interpolation entsprechende Darstellung der Eliminations-Resultante. *Journal für die Reine und angewandte Mathematik*, 1860(57):111–121, 1860.

[BR89] P. Baldi and Y. Rinott. On normal approximations of distributions in terms of dependency graphs. *Ann. Probab.*, 17(4):1646–1650, 1989.

[BR10] R. N. Bhattacharya and R. R. Rao. *Normal Approximation and Asymptotic Expansions*, volume 64 of *Classics in Applied Mathematics*. SIAM, 2010.

[Buf12] A. I. Bufetov. A central limit theorem for extremal characters of the infinite symmetric group. *Functional Analysis and Its Applications*, 46(2):83–93, 2012.

[Cha12] S. Chatterjee. The missing log in large deviations for triangle counts. *Random Structures & Algorithms*, 40(4):437–451, 2012.

[CV11] S. Chatterjee and S.R.S. Varadhan. The large deviation principle for the Erdős-Rényi random graph. *Eur. J. Comb.*, 32(7):1000–1017, 2011.

[DE12] H. Döring and P. Eichelsbacher. Moderate deviations via cumulants. To appear in *Journal of Theor. Probability*, 2012.

[DE13] H. Döring and P. Eichelsbacher. Moderate deviations for the determinant of wigner matrices. `arXiv:1301:2915v2 [math.PR]`, 2013.

[DKN11] F. Delbaen, E. Kowalski, and A. Nikeghbali. Mod-ϕ convergence. `arXiv:1107.5657v2 [math.PR]`, 2011.

[DZ98] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, 2nd edition, 1998.

[EK40] P. Erdős and M. Kac. The Gaussian law of errors in the theory of additive number theoretic functions. *American Journal of Mathematics*, 62:738–742, 1940.

[ER60] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17–61, 1960.

[Ess45] C.-G. Esseen. Fourier analysis of distribution functions. A mathematical study of the Laplace-Gauss law. *Acta Math.*, 77(1):1–125, 1945.

[Fel71] W. Feller. *An Introduction to Probability Theory and Its Applications, Volume II.* Wiley Series in Probability and Mathematical Statistics. Wiley, 2nd edition, 1971.

[FM12] V. Féray and P.-L. Méliot. Asymptotics of q-Plancherel measures. *Probab. Th. Rel. Fields*, 152(3-4):589–624, 2012.

[FO90] P. Flajolet and A.M. Odlyzko. Singularity analysis of generating functions. *SIAM J. Discrete Math.*, 3:216–240, 1990.

[FS90] P. Flajolet and M. Soria. Gaussian limiting distributions for the number of components in combinatorial structures. *Journal of Combinatorial Theory. Series A*, 53:165–182, 1990.

[FS09] P. Flajolet and R. Sedgewick. *Analytic Combinatorics.* Cambridge University Press, 2009.

[HKO01] C.P. Hughes, J.P. Keating, and N. O'Connell. On the characteristic polynomial of a random unitary matrix. *Comm. Math. Phys.*, 220(2):429–451, 2001.

[Hör94] L. Hörmander. *Notions of convexity*, volume 127 of *Progress in Mathematics*. Birkhäuser, 1994.

[Hwa94] H.-K. Hwang. *Théoremes limites pour les structures combinatoires et les fonctions arithmétiques, phd thesis, Ecole Polytechnique*. PhD thesis, Ecole Polytechnique, 1994.

[Hwa96] H.-K. Hwang. Large deviations for combinatorial distributions. I. Central limit theorems. *Ann. Appl. Probab.*, 6(1):297–319, 1996.

[IK99] V. Ivanov and S. Kerov. The algebra of conjugacy classes in symmetric groups, and partial permutations. In *Representation Theory, Dynamical Systems, Combinatorial and Algorithmical Methods III*, volume 256 of *Zapiski Nauchnyh Seminarov POMI*, pages 95–120, 1999.

[Ilt95] M. Iltis. Sharp asymptotics of large deviations in $\mathbb{R}^d$. *J. Theoretical Prob.*, 8:501–524, 1995.

[IO02] V. Ivanov and G. Olshanski. Kerov's central limit theorem for the Plancherel measure on Young diagrams. In *Symmetric Functions 2001*, volume 74 of *NATO Science Series II. Mathematics, Physics and Chemistry*, pages 93–151, 2002.

[IV07] I.Bárány and V. Vu. Central limit theorems for gaussian polytopes. *Ann. Probab.*, 35(4):1593–1621, 2007.

[Jan88] S. Janson. Normal convergence by higher semiinvariants with applications to sums of dependent random variables and random graphs. *Ann. Probab.*, 16(1):305–312, 1988.

[Jan04] S. Janson. Large deviations for sums of partly dependent random variables. *Random Structures & Algorithms*, 24(3):234–248, 2004.

[JKN11] J. Jacod, E. Kowalski, and A. Nikeghbali. Mod-Gaussian convergence: new limit theorems in probability and number theory. *Forum Mathematicum*, 23(4):835–873, 2011.

[JŁR00] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*, volume 45 of *Wiley Series in Discrete Mathematics and Optimization*. Wiley-Interscience, 2000.

[KN10] E. Kowalski and A. Nikeghbali. Mod-Poisson convergence in probability and number theory. *Intern. Math. Res. Not.*, 18:3549–3587, 2010.

[KN12] E. Kowalski and A. Nikeghbali. Mod-Gaussiam distribution and the value distribution of $\zeta(1/2+it)$ and related quantities. *J. London Math. Soc.*, 86(2):291–319, 2012.

[KNN10] E. Kowalski, J. Najnudel, and A. Nikeghbali. A characterization of limiting functions arising in mod-* convergence. Preprint, 2010.

[KS00a] J. P. Keating and N. C. Snaith. Random matrix theory and L-functions at $s = 1/2$. *Comm. Math. Phys.*, 214(1):91–110, 2000.

[KS00b] J. P. Keating and N. C. Snaith. Random matrix theory and $\zeta(1/2 + \mathrm{i}t)$. *Comm. Math. Phys.*, 214(1):57–89, 2000.

[Kub72] J. P. Kubilius. On large deviations of additive arithmetic functions. *Trudy Mat. Inst. Steklov.*, 128:163–171, 1972.

[KV81] S. V. Kerov and A. M. Vershik. Asymptotic theory of characters of the symmetric group. *Funct. Anal. Appl.*, 15(4):246–255, 1981.

[LS59] V.P. Leonov and A.N. Shiryaev. On a method of calculation of semi-invariants. *Theory Prob. Appl.*, 4:319–329, 1959.

[Mél12] P.-L. Méliot. Fluctuations of central measures on partitions. In *Proceedings of the 24th International Conference on Formal Power Series and Algebraic Combinatorics (Nagoya, Japan)*, pages 387–398, 2012.

[Ney83] P. Ney. Dominating points and the asymptotics of large deviations for random walk on $\mathbb{R}^d$. *Ann. Probab.*, 11:158–167, 1983.

[NZ13] A. Nikeghbali and D. Zeindler. The generalized weighted probability measure on the symmetric group and the asymptotic behavior of the cycles. *Annales de l'Institut Henri Poincaré, to appear*, 2013.

[Oko00] A. Okounkov. Random matrices and random permutations. *Internat. Math. Res. Notices*, 20:1043–1095, 2000.

[Pet75] V. V. Petrov. *Sums of independent random variables*, volume 82 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, 1975.

[Rad09] M. Radziwill. On large deviations of additive functions. B.Sc thesis, `arXiv:09095274v4 [math.NT]`, 2009.

[RSS78] R. Rudzkis, L. Saulis, and V. A. Statulevičius. A general lemma for large deviations. *Lithuanian Math. J.*, 18:99–116, 1978.

[Ruc88] A. Ruciński. When are small subgraphs of a random graph normally distributed? *Probab. Th. Rel. Fields*, 78(1):1–10, 1988.

[Sch93] R. Schneider. *Convex Bodies: The Brunn-Minkowski Theory*, volume 44 of *Encyclopaedia of Mathematics and Its Applications*. Cambridge University Press, 1993.

[Śni06a] P. Śniady. Asymptotics of characters of symmetric groups, genus expansion and free probability. *Discrete Math.*, 306(7):624–665, 2006.

[Śni06b] P. Śniady. Gaussian fluctuations of characters of symmetric groups and of Young diagrams. *Probab. Th. Rel. Fields*, 136(2):263–297, 2006.

[SS91] L. Saulis and V. A. Statulevičius. *Limit Theorems for Large Deviations*. Kluwer Academic Publications, 1991.

[Ste93] E. M. Stein. *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, 1993.

[Ten95] G. Tenenbaum. *Introduction to analytic and probabilistic number theory*, volume 46 of *Cambridge studies in advanced mathematics*. Cambridge University Press, 1995.

[Tho64] E. Thoma. Die unzerlegbaren, positive-definiten Klassenfunktionen der abzählbar unendlichen symmetrischen Gruppe. *Math. Zeitschrift*, 85:40–61, 1964.

LaBRI, Université Bordeaux 1 — 351 cours de la Libération, 33400 Talence, France
E-mail address, V. Féray: `feray@labri.fr`

Institut für Mathematik, Universität Zürich — Winterthurerstrasse 190, 8057 Zürich, Switzerland
E-mail address, P.-L. Méliot: `pierre-loic.meliot@math.uzh.ch`

Institut für Mathematik, Universität Zürich — Winterthurerstrasse 190, 8057 Zürich, Switzerland
E-mail address, A. Nikeghbali: `ashkan.nikeghbali@math.uzh.ch`